Technology Generations
handling complex User Interfaces

CIP-DATA LIBRARY TECHNISCHE UNIVERSITEIT EINDHOVEN

Docampo Rama, M.
Technology generations handling complex user interfaces / by M. Docampo Rama. - Eindhoven : Technische Universiteit Eindhoven, 2001. - Proefschrift.
ISBN 90-386-0913-2
NUGI 719
Subject Headings: Generations ; User interfaces ; Cognition ; Ergonomics

druk: Universitaire Drukkerij, Eindhoven
omslag: Philips Design

The work in this thesis has been carried out under the auspices of the research school J.F. Schouten Instituut for User-System Interaction Research.

Technology Generations handling complex User Interfaces

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de Rector Magnificus, prof.dr. M. Rem, voor een commissie aangewezen door het College voor Promoties in het openbaar te verdedigen op donderdag 29 maart 2001 om 16.00 uur

door

Maria de los Milagros Docampo Rama

geboren te Tegelen

Dit proefschrift is goedgekeurd
door de promotoren:

prof.dr. H. Bouma
en
prof.dr. D.G. Bouwhuis

Copromotor:
prof.dr. H. de Ridder

Preface

The research presented in this thesis is part of the project Technology Generations, commissioned by Philips Design in 1996 and partially subsidised by the Ministry of Economical Affairs (EZ). The emphasis of this project was to analyse age and generation-related differences in attitude towards technology, and purchase and usage of technology. The project was initially coordinated by Cees de Bont, who was over a five year period followed by Sidney Vastenhout, Hans de Waard and Carel Vereijken, and concluded by the author of this thesis.

The project involved four teams. The first team focused on generation related attitude towards design of technology and was staffed by Liesbeth Scholten, Carel Vereijken, and Juliana Kelly. The second team researched the sociological aspects of technology generations and consisted of Anne-Geerte van de Goor and Henk Becker. Their book on ‘Technology Generations in the Netherlands: A sociological analysis’ has been published in January 2001. The third team analysed the history of technology and the attitude of consumers towards technology and consisted of Frederiek van der Kaaden, Geert Verbong, and Harry Lintsen. The results are expected to be published at the end of 2001. The fourth group dealt with age and generation factors that influence the handling of complex user interfaces and consisted of Herman Bouma, Don Bouwhuis, Huib de Ridder, Jan Rietsema, Boris de Ruyter, John de Vet and the author of this thesis. The results are published in this thesis. I would like to thank all team members for their contribution to the project. In particular, I would like to thank Anne-Geerte van de Goor, Frederiek van der Kaaden, and Liesbeth Scholten for their enthusiastic cooperation and support. It was fun! Furthermore, I would like to thank the designers Femke Bourgogne, Kirsten Moelker and Paul Manders from Philips Design for designing the cover of this thesis.

I would like to express my gratitude to Anne-Sophie Melenhorst, my office mate and discussion partner on many issues, and Willem Paul Brinkman, who supported me on the modelling of the automation tool outlined in Chapter 5. Special thanks go to Maxim Schram, who implemented the automation tool and supported me with the empirical studies in Chapter 5. Moreover, his great Artifact-

music kept me going in the evening hours. I am also indebted to Arlette Oosterholt and Luc Gijs for their assistance with the empirical studies in Chapter 3 and 4, Gerard Hollemans and Esther Klabbbers for carefully reading drafts of this thesis, Mrs. Geven, and Mrs. Willems, and Mrs. Ruyten for providing me with volunteers for the studies described in this thesis, and in particular to all 210 volunteers. This thesis would not have been completed without them. I also thank Pyns Bosker from ScreenWise and Zjev van Dun from CINOP, who gave me the opportunity to discuss and apply methods presented in this thesis in real life.

Early 1999 we started communicating the Technology Generations results inside Philips. By now, the concept of Technology Generations has been proven to be very valuable within trend research, e-business and services for strategic projects. Therefore, I acknowledge our customers Philips Consumer Electronics and Philips Semiconductors, who embraced this new research and collaborated with us to make it applicable for future concepts and up-stream services.

Furthermore I thank my colleagues at Philips Design with whom I developed Technology Generation-related tools that support our client's strategic processes and our approach to Strategic Design.

Special acknowledgments are due to my colleagues at IPO Center for User-System Interaction, who acted as research partners and friends. This made IPO a tremendously inspiring working environment.

Finally, I would like to thank my parents, brother, and family in-law, who supported me in every way they could. I also would like thank my husband Marco Verhoeven, who was my sparring partner in technical and emotional issues surrounding this thesis. Having someone to share your life with makes writing a thesis seem insignificant.

para mi querido marido Marco

Contents

1

Introduction

1.1 Problem statement

The technology revolution of the twentieth century has dramatically changed the teacher-pupil relation of older and younger people into an ambiguous or even reverse relation. For centuries older adults have been the obvious teachers of young persons on behaviour, social skills and profession. Now it may be often observed that younger persons teach older adults where new technology is involved.

Did you ever notice how surprised parents are to see their children use a mouse-pointing device quickly, play a computer game with ease, or program a video recorder? At such moments, parents may reflect upon the difficulties they had learning it. It is also not unusual to see grandchildren teach their grandparents how to use a video recorder, mobile phone, computer, or Internet. And recall the commotion that arises when the operator of the cable TV system changes channel frequencies. Studies on the ease of use of information, communication, and entertainment devices abbreviated as ICE confirm that the majority of adults have trouble using them [Freudenthal, A., 1998; Czaja and Sharit, 1998; Rogers and Fisk, 2000]. However, Freudenthal, A. [1998] found that also little children will fail to program a television if they have only experienced computer games.

The past century has been a period full of revolutionary technological innovations entering society on a massive scale. In the beginning of the twentieth century the telephone was introduced, followed by the radio, gramophone, refrigerator, and vacuum cleaner among other electrical devices. Later on, especially the number and variety of ICE devices grew explosively. The television, video recorder, video camera, computer, compact disc, remote control devices, wireless phone, and mobile phone increased a person's information and communication radius and extended entertainment possibilities. Mutual interaction between the technological innovations and the consumer has substantially changed the way of life at home and at work [Schot, Lintsen, Rip and De la Bruhéze, 2000].

As technological inventions involving consumer products continue to spread in the 21^{st} century it is important for adults to keep track of them. Only then they have the option to use a device whenever they need or want to. The fact that in particular older people seem to lose track of novel technology has practical implications for the industry. Industry is committed to market technology in service of society, which implies development of useable products for everyone, including older adults. Moreover, society is greying fast, and many older adults of today and tomorrow have much more purchasing power than older people used to have. Neglecting the older adults' problems in the use of technology also means a missed opportunity for industry to grow its market share and revenue.

This thesis is concerned with the question why consumer electronics products are felt to be difficult to use by older adults. The focus will be on human capabilities and limitations versus simple and complex aspects of the systems' user interfaces.

1.1.1 *User interface complexity*

Older persons often mention that devices of the early days were much easier to use than current consumer electronics products. If so, could it mean that today's user interfaces are complex for everybody, including younger adults?

To answer this question we first need to know more about the differences between today's user interfaces and those of decades ago. Although most human-computer studies start their introduction with a phrase stressing the complexity of current devices, there has not been any systematic study that analysed the historical development of user interfaces for consumer electronics over the past century, to the best of our knowledge.

Obviously, the technological developments have influenced the way user interfaces evolved over the years. A typical user interface of the first electrical devices contains some robust mechanical push buttons, sliders, and switches that activate basic functions of the device. Nowadays, this does not often, or not at all, come across on user interfaces. In this thesis, a historical analysis will be presented to study how user interfaces of consumer electronics products changed over the years. This will enable a meaningful discussion on the change in complexity of the user interface.

This thesis is concerned with the analysis of user interface modifications that affected the ease of use for everybody, and in particular for older adults.

1.1.2 *Ageing*

Although present-day user interfaces may be difficult to use, it is remarkable that mainly older adults complain about having difficulties. This may suggest that age-related factors influence the user's handling behaviour. User-system interaction studies confirm that the difficulties experienced using a system increase with age. Older adults require more actions than younger people when using an interactive system to obtain the same result [Freudenthal, T.D., 1998], are less effective using a computer [Kelley, 1996], need more time to finish computer tasks [Salthouse, Hambrick, Lukas and Dell, 1996], and commit more procedural errors during data entry tasks [Czaja and Sharit, 1998]. Which age-related cognitive changes affect the older user's performance?

Dealing with a device can be facilitated by knowledge. Anderson [1983] distinguishes two types of knowledge, procedural and declarative knowledge. Procedural knowledge is stored information about how to do things. It includes motor and cognitive skills. This information cannot easily be verbalized. Declarative knowledge contains stored 'what' information about the environment. This type of knowledge also contains information about how to do things i.e., what steps have to be taken. According to Anderson, learning in this context is the process of qualitative changes from declarative to procedural knowledge.

Cognitive skills are needed to acquire well-established procedural knowledge [Anderson, 1983; Anderson, 1990]. For example, a good reasoning ability is needed to interpret which information is relevant and what the logic is behind the procedures. Furthermore, the user's working memory is of importance to keep relevant declarative knowledge available during the execution of tasks. Working memory is a temporary storage in which incoming information, acquired from the environment and from the person's long term memory, is kept available for the completion of the task [Baddeley, 1986].

It is known that most cognitive abilities change with age. Although there are many individual differences in ability changes, it can be said that roughly, vocabulary and factual knowledge increase with age. Sensory memory and procedural memory for skills learned a long time ago do not change that much with age [Craik and Bosman, 1992], but most other cognitive abilities show a decline with age [Salthouse, 1991; Craik and Bosman, 1992]. Studies of, for example, Salthouse, Hambrick, Lukas and Dell [1996], Czaja and Sharit [1998], and Freudenthal, T.D. [1998] provided evidence that cognitive factors like processing speed, working memory, and reasoning ability relate to decreasing performance with age while handling tasks on complex systems. Older users also seem to apply a different user strategy than younger users. Welford [1958] found that the speed-accuracy tradeoff changes with increasing age from a speedy and inaccurate way of playing around (younger adults) to a cautious but more accurate performance (older adults).

In line with these studies, next to the complexity of user interfaces this thesis is concerned with the role of cognitive abilities that may decline with age in the user's performance on controlling current consumer electronics products.

1.1.3 *Technology generations*

The user's learning ability in new situations can be facilitated by reusing knowledge of earlier but related situations. As the relevant knowledge base of each person depends on the type of user interfaces he/she worked with, it is possible that older people have built up a different knowledge base than younger people. However, not all knowledge is expected to be that useful, as present-day devices are different from those of decades ago. This means that part of the knowledge acquired in the past has to be suppressed. But how easy is it to forget it?

People often rely on their most established knowledge, which is acquired during their adolescence [Rubin, Rahhal and Poon, 1998]. This is because the period of adolescence and young adulthood has been found to be the critical period in which experience and cognitive abilities influence each other for further maturation [Sroufe and Cooper, 1988]. But not only developmental studies stress this period. Also sociological studies (e.g. Glenn [1974] and Becker [1992]) on cohort analyses indicate that the life period between 10 and 25 years is a most critical period in which people acquire norms, values, and skills that influence future behaviour. Therefore, this period is also called the formative period. After this period adults may become less susceptible to attitudinal and behavioural changes.

A *generation* has been defined as a range of birth cohorts that display similar behaviour, and share norms or values that are based on common sociological environments during their formative years. Note that a generation can only be identified if the environment during the formative period is stable long enough to affect the behaviour of many subsequent birth cohorts [Jansen and Van den Wittenboer, 1992]. In this way Becker [2000] distinguished five generations based on the differences in socio-historical developments during the formative period separated by discontinuous events, such as wars and economic depressions: the prewar generation (1910-1920), silent generation (1930-1945), baby-boom generation (1946-1954), lost generation (1955-1969), and generation X (born in 1970 or later).

Sackmann and Weymann [1994] investigated whether generations can be distinguished from subsequent birth cohorts that currently display similar behaviour towards technology based on macro-technological events in their formative period. They were able to identify four *technology generations* based on the types of technology that were available in the formative periods and the types of products they currently possess: the early technical generation (born before 1938), household revolution generation (born between 1939 and 1948), the increasing technological household generation (born between 1949 and 1963), and the computer generation (born in 1964 or later).

User interface developments imply that users need to constantly adapt to new user interfaces and learn to operate them; after their formative period too. Therefore, it might be the case that technology generations can be distinguished based on the type of user interface experience during the formative period. Technology generations that did not grow up with present-day user interfaces are expected to have most difficulties dealing with present-day devices. They need to learn to understand new procedures of present-day user interfaces, which might be much more difficult to manage after the formative period.

This leads to another concern of the thesis, next to those mentioned in earlier sections, namely whether the user's performance in dealing with current consumer products is related to generation differences based on user interface experience.

1.2 Objectives

The objective of this thesis is to investigate which aspects of current user interfaces affect the handling by users of different age groups. Performance is further analysed with respect to age and technology generation effects by studying the role of cognitive abilities and generation-specific experience with user interfaces.

Chapter 2 describes a historical analysis that maps out the user interface developments of consumer electronics products over the past century. Based on this analysis we will try to distinguish several generations of subsequent birth cohorts that experienced a similar type of user interface in their formative years. Chapter 2 concludes with a study in which we empirically analyse the technology generation boundaries.

Chapter 3 investigates by means of laboratory studies whether the user's error performance and task duration are affected by age, technology generation, and user interface complexity. However, it is not easy to distinguish an age effect from a generation effect as both are related to the user's year of birth. Sociologists developed a method to distinguish an age effect from a cohort effect [Glenn, 1977]. Chapter 3 first explains this method and describes how it is applied in this thesis. Subsequently, two studies will be described and discussed. The first consists of an experiment in which four age groups between 25 and 75 year of age are confronted with simulated devices that differ in one of the user interface features that changed over the past century. Age and technology generation effects have been tested by analysing the error and task duration performance change over trials and across age groups. The second study consists of a similar design to analyse whether the results of the first study can be generalized towards more complex devices with different functionality.

Chapter 4 investigates how error performance and task duration across and within technology generations, as found in chapter 3, are related to visuo-spatial working memory, reasoning ability and a possible speed-accuracy tradeoff. For this, the second study of chapter 3 was repeated for the age range of 25 to 80 years of age.

Chapter 5 studies user strategies within the oldest technology generation when using a simulation of a realistic consumer electronics product. The strategies are compared with those of a younger technology generation that is known to have grown up with such user interfaces. Five age groups between 20 and 80 years were compared. The user's strategy is investigated by analysing the drop-out rate, speed-accuracy trade-off, and types of redundant actions. Three automated tools were used for this, enabling the analysis of larger user groups.

Chapter 6 reflects on the conclusions drawn in the prior chapters and discusses the implications of the results for user interface design, instruction of use, and further research.

As some chapters have been submitted as articles to international journals and conferences, and others will be submitted soon, part of the information will be repeated across the chapters to introduce the new reader to relevant concepts and prior results.

2

Characterization of technology generations on the basis of user interfaces

Technology generations can be defined as generations of people that differ in the type of technology experience in their formative years (i.e. before the age of 25). The formative experience of these technology generations may have consequences for their current behaviour towards present-day Information, Communication, and Entertainment appliances (ICE devices). This chapter focuses on the user interface experience before the age of 25 of the various generations. We investigate how the user interface of ICE devices has evolved over the years, and what consequences it has for the concept of technology generations. Results show that a major user interface-change occurred in the beginning of the eighties. Two technology generations have been identified: the electro-mechanical generation (born before 1960) and the display generation (born in 1960 and later). [1]

---

[1] Docampo Rama and Van der Kaaden [2001]

2.1 Introduction

Older users often mention having difficulties using present-day Information, Communication, and Entertainment appliances (abbreviated in this study as ICE appliances), such as the mobile phone, the video recorder, and the television [Stewart, 1992; Freudenthal, A., 1998]. These problems may be due to changes in cognitive abilities with age (e.g. Salthouse [1991], Czaja and Sharit [1998]), lack of prior experience with similar devices, and the increasing complexity of present-day user interfaces. This study will focus on the influence of the last two aspects; the experience with and the complexity of present-day ICE devices.

When a person learns to use a new user interface he/she often relies on former user-interface experience [Singley and Anderson, 1985; Searleman and Douglas, 1994]. The mismatch between the knowledge gained during prior experience with similar devices and the knowledge required to use the new one makes it likely that users will encounter difficulties [Singley and Anderson, 1985; Kelley, 1996]. Difficulties among older users could be related to their lack of knowledge about present-day user interfaces [Kelley, 1996]. Moreover, older users grew up with totally different user interfaces, and may therefore have acquired knowledge that is not applicable to present-day user interfaces.

As the user interface of ICE devices has changed significantly over the years, not all user interface experience in the past will be useful when using present-day ICE devices. This means that part of the knowledge acquired in the past has to be suppressed. But how easy is it to forget past experience?

Studies concerning cohort or generation analyses [Glenn, 1974; Cutler and Kaufman, 1975] have indicated that people are likely to become less susceptible to change in attitudes, norms, values, and behaviour after young adulthood. Among the reasons given for this attitudinal and behavioural rigidity are neurological change, decreased social and geographic mobility, and changes in social contacts.

The period of young adulthood and adolescence, operationalized as the period between 10 and 25, is also known as the formative period and plays a key role in the formation of these attitudes and behaviours. In this period, people have to master a number of crucial transitions, for example from school to work, and from parental home to living independently. During these events, people acquire important values, norms, and skills that influence the way they behave during the rest of their life [Becker, 1992].

Child development studies indicate that a person's cognitive abilities, such as hypothetico-deductive reasoning, reach their highest level during adolescence (e.g. Piaget and Inhelder [1973], Keating [1980], and Neimark [1982]). In this period the person's social relationships, (technical) hobbies, and other life experiences both influence and are influenced by cognitive advances [Sroufe and Cooper, 1988]. Consequently, the formative period is the life phase in which complex systems and their user interfaces can be optimally understood and learned to use in all its facets. As a result, technology experience in the formative period may have a great impact on the user's way of handling technology later in life.

If the formative period of several birth-cohorts is influenced by the same sociological macro-events during a substantial period of time, a generation of subsequent cohorts can be distinguished that displays similar behaviour today [Glenn, 1974]. Empirically, this generation effect is difficult to distinguish from an age effect, as both generation and age are measured by the subject's year of birth. Sociological methods tackle this problem by analysing the behaviour trend across age groups. A continuous trend suggests an age effect, whereas a discontinuous trend indicates a generation effect [Glenn, 1977].

Sackmann and Weymann [1994] have shown that generations based on similar technology experience in the formative period can be identified. People that experienced the availability of the same types of domestic appliance during the formative years, display similar usage behaviour. Such a group has been defined as a technology generation. Also Van de Goor [1998] and Van de Goor and Becker [2001] have shown that technology generations differ in current purchase and usage behaviour. These results suggest that generation-specific experience may be an important explanatory factor for age group differences in current behaviour.

This study focuses on user interface experience of present-day ICE appliances in the formative period to distinguish different technology generations. The generation-specific experience will be based on the evolution of user interfaces and their diffusion time. For this purpose two studies were conducted. Firstly, a historical analysis was made of user interface changes of three appliances, namely television, video recorder, and telephone. The outcomes have been used to define technology generations based on user interface changes. Secondly, a survey was carried out to verify the boundaries of the generations.

Section 2.2 describes the historical analysis of the user interface changes and discusses the generation boundaries that can be derived from it. It is followed in section 2.3 by a verification of the proposed boundaries via a survey. Main conclusions are mentioned in section 2.4.

2.2 Historical developments of user interfaces in the Netherlands

2.2.1 *Introduction*

User interfaces can be described in a structured and detailed way by specifying their interaction style. De Vet and De Ruyter [1996] proposed that an interaction

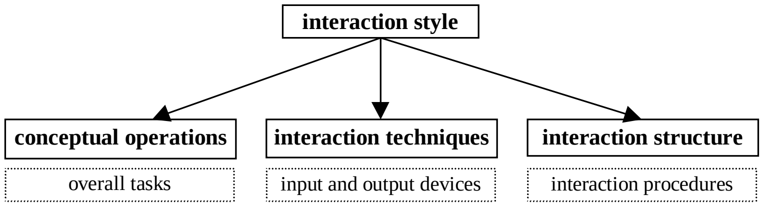


Figure 2.1: Decomposition of the interaction style into conceptual operations, interaction techniques, and structure [De Vet and De Ruyter, 1996]. The dotted rectangles underneath explain the user's view per interaction style component.

style can be decomposed into generic user-task options or functionality (conceptual operations), organised according to a specific operational procedure (interaction structure), and using specific input and output devices (interaction techniques). Figure 2.1 shows this framework, which can also be used for historical analyses of user interface changes of ICE appliances over the years.

Specification of interaction style components of ICE user interfaces over the years will lead towards differentiation of interaction styles, which are based on the periods in which all three interaction style components (conceptual operations, interaction techniques and structure) have changed.

2.2.2 *Method*

Material and procedure

The analyses were restricted to user interface changes of the television, telephone, and video recorder. The Dutch consumer's magazine [Consumentenbond, 1953 to 1997] and technical manuals were used for the historical overview of the functionality, input and output devices, and procedural aspects of the ICE devices in question. Figures mentioned in the literature were used to add the year of introduction of each product to the overview.

In the text of the magazine we searched for comments concerning the functionality, input and output devices, and procedures available on the ICE devices under consideration. Also pictures were used to search for these aspects. The technical manuals were used to deduce the procedures in more detail. Each time a new interaction style component was observed, it was registered in the historical overview.

2.2.3 *Results*

The television

The black and white television entered the Dutch market in 1951 after some experimental try-outs from 1936 on [Baudet, 1986]. Figure 2.2 depicts the most prominent conceptual operations, interaction techniques, and structure changes over the years. Certain features listed in the figure were only available on the market for a short period.

During the fifties and sixties, the user interface of the television comprised only basic functionality, mechanical interaction techniques, and a straightforward interaction structure that related interaction techniques directly to the conceptual operations. In the seventies, the number of features implemented on the television increased and new interaction techniques were introduced, such as the wireless remote control, touch buttons, and display. These interaction techniques became more prominent and replaced most mechanical attributes of the television in the eighties. No interaction structure changes were introduced in the seventies, but in the eighties totally new procedures for programming were implemented next to a large extension of functionality and software-based interaction techniques. The remote control that was introduced appeared to be the first device separating the user interface from the appliance. The one-to-one relation between the interaction techniques and the conceptual operations disappeared, and most interaction techniques had more functions depending on their mode. Changes resulted in fewer buttons on the consumer product itself and many buttons on the remote control. In the nineties, new elements were added to all interaction style components: specialized functionality came onto the market that was focused on services to the user. Visual, software-based, interaction techniques replaced most hardware-based attributes, and a new interaction structure, derived from the computer and known as a menu structure, was introduced, for which navigation and selection procedures have to be used. Beside that, automatic programming procedures were developed to reduce the complexity of the interaction structure.

The telephone

Similar trends as observed for the television were found for the telephone, which came onto the Dutch market in 1877 [Baudet, 1986]. There are two exceptions in the evolution of the telephone user interface: in the beginning, the operator functioned as the user interface of the telephone after picking up the phone and activating it by a swing. Around 1930 the user interface changed towards a device with a rotary dial that enabled calling without having to ask the operator to make the connection [Tours, 1981]. Secondly, the expansion of functionality and the change of interaction techniques as seen on the television in the seventies were not seen on the telephone until the eighties. See figure 2.3 for an extensive overview of the changes. Note that most changes occurred in the nineties. Certain features listed in the table were only available on the market for a short period.

The video recorder

Also the consumer video recorder, on the Dutch market at the end of the seventies, showed a similar trend as the television and the telephone. At the end of the seventies the consumer video recorders contained some basic features. Specialized functionality was gradually introduced in the eighties. See figure 2.4 for an extensive overview of these user interface changes. Certain features listed in the table were only available on the market for a short period.

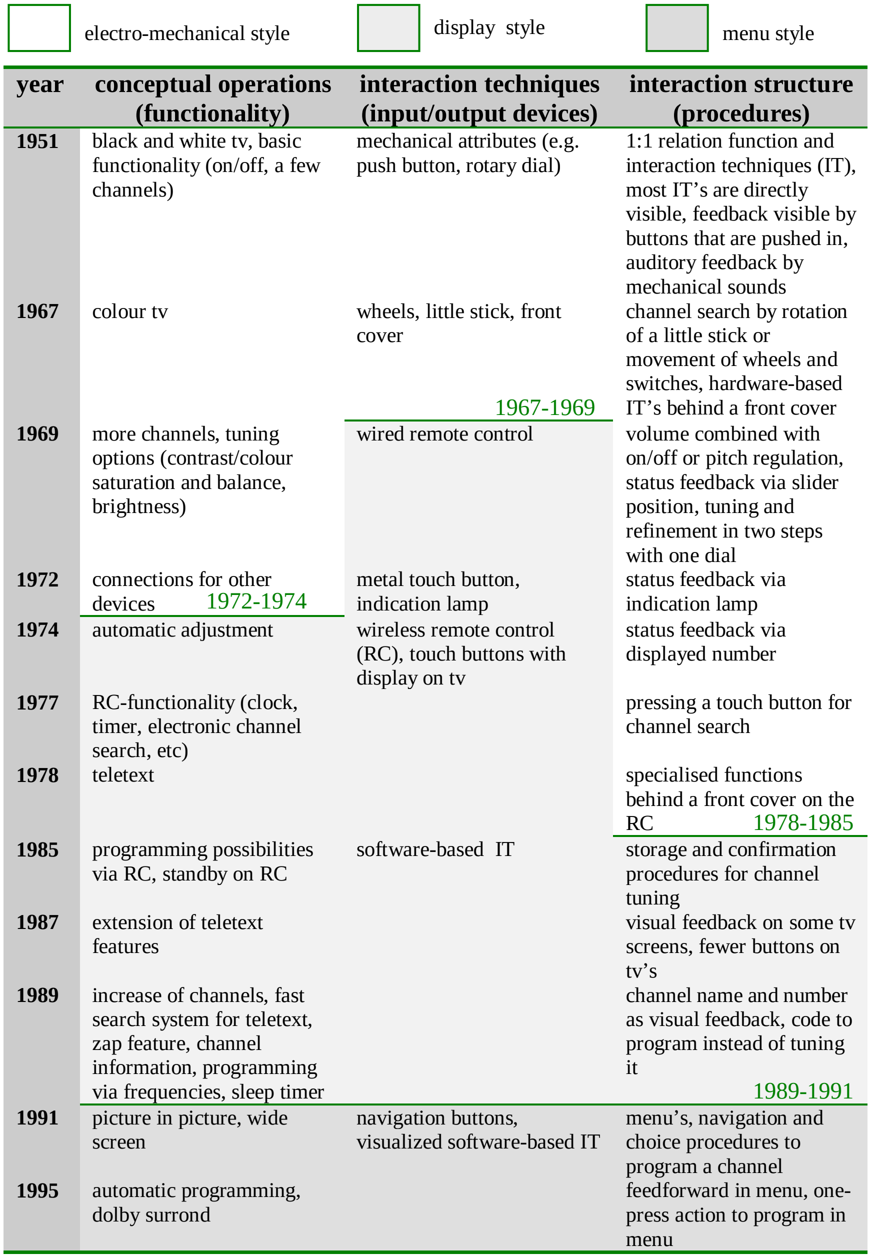
2.2.4 *Discussion*

The purpose in this study is to distinguish technology generations based on the user interface experience in their formative period. We will focus first on the historical analysis described in the previous section, which guides us towards the different interaction styles over the years. The historical analysis does not indicate when the majority of the population has experienced it. For this we need to know more about its diffusion time. By estimating when the majority of the people will have experienced a certain interaction style in their formative period, we can arrive at distinguishable technology generations based on user interfaces. These aspects will be discussed in this section. For more details, see Docampo Rama and Van der Kaaden [1998].

Interaction styles

Part of the user interface changes on the telephone, television, and video recorder occurred simultaneously. Based on the period in which all interaction style components of one device changed, three interaction styles can be distinguished (see the borders drawn in the figures 2.2, 2.3, and 2.4).

In the first interaction style implemented on the television, telephone and video recorder the user executed by hardware-based electro-mechanical interaction techniques, containing only basic functionality, directly visible on the user interface with almost always a one-to-one relation between interaction techniques and functionality. In this thesis, this style will be named the electro-mechanical style. The electro-mechanical style was first implemented on the telephone in the 1930’s [Tours, 1981], but also on the radio [Baudet, 1986]. Later on this style was also applied to the television (since 1964). The video recorder was implemented in this style as well, with exception of the huge amount of functionality in video recorders for professional use. Around 1979 the video recorder came on the market



□ electro-mechanical style ░ display style ▒ menu style

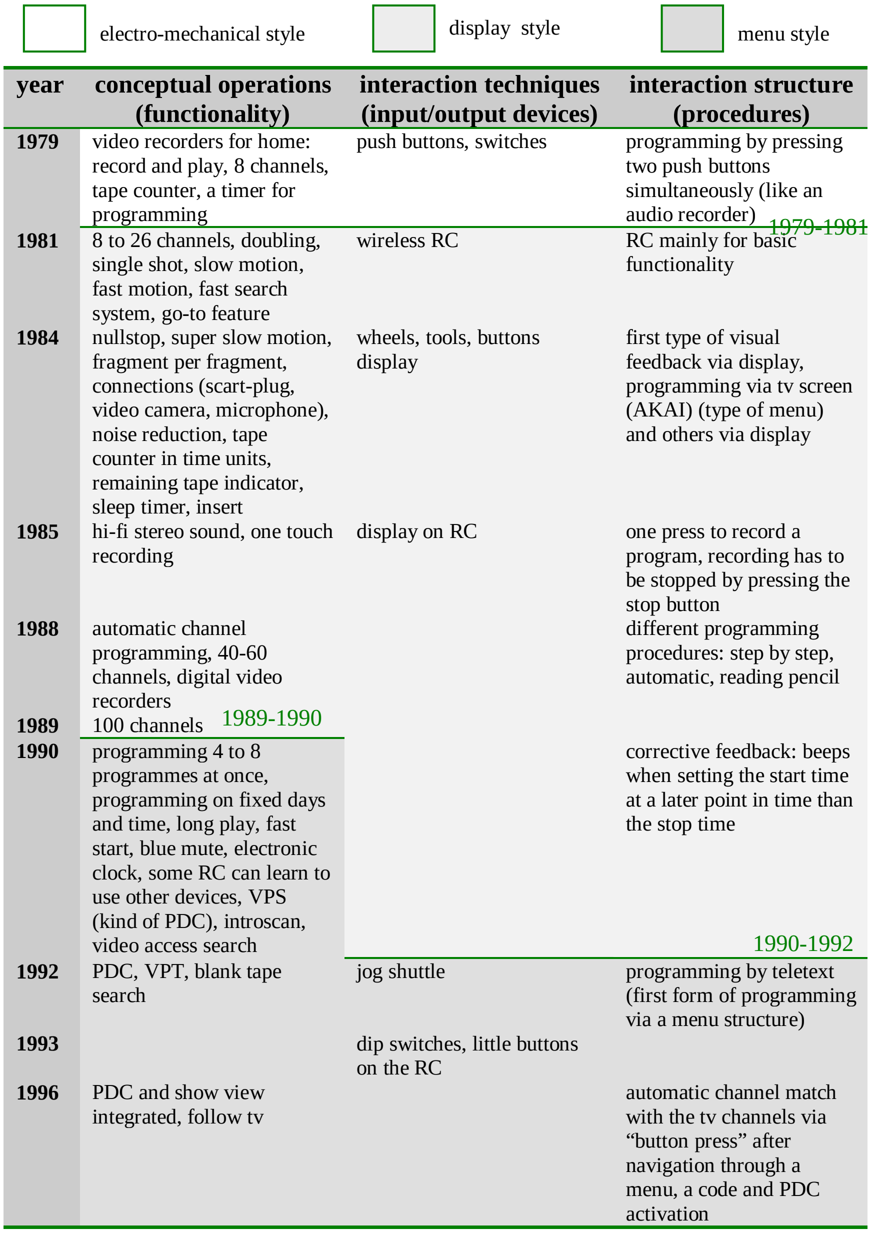
year	conceptual operations (functionality)	interaction techniques (input/output devices)	interaction structure (procedures)
1951	black and white tv, basic functionality (on/off, a few channels)	mechanical attributes (e.g. push button, rotary dial)	1:1 relation function and interaction techniques (IT), most IT's are directly visible, feedback visible by buttons that are pushed in, auditory feedback by mechanical sounds
1967	colour tv	wheels, little stick, front cover 1967-1969	channel search by rotation of a little stick or movement of wheels and switches, hardware-based IT's behind a front cover
1969	more channels, tuning options (contrast/colour saturation and balance, brightness)	wired remote control	volume combined with on/off or pitch regulation, status feedback via slider position, tuning and refinement in two steps with one dial
1972	connections for other devices 1972-1974	metal touch button, indication lamp	status feedback via indication lamp
1974	automatic adjustment	wireless remote control (RC), touch buttons with display on tv	status feedback via displayed number
1977	RC-functionality (clock, timer, electronic channel search, etc)		pressing a touch button for channel search
1978	teletext		specialised functions behind a front cover on the RC 1978-1985
1985	programming possibilities via RC, standby on RC	software-based IT	storage and confirmation procedures for channel tuning
1987	extension of teletext features		visual feedback on some tv screens, fewer buttons on tv's
1989	increase of channels, fast search system for teletext, zap feature, channel information, programming via frequencies, sleep timer		channel name and number as visual feedback, code to program instead of tuning it 1989-1991
1991	picture in picture, wide screen	navigation buttons, visualized software-based IT	menu's, navigation and choice procedures to program a channel
1995	automatic programming, dolby surrond		feedforward in menu, one-press action to program in menu

Figure 2.2: Overview of user interface changes on the television from 1951-1996. The years within the figure indicate the transition period between new styles.

☐ electro-mechanical style ▨ display style ▩ menu style

year	conceptual operations (functionality)	interaction techniques (input/output devices)	interaction structure (procedures)
1930	call	rotary dial	1:1 relation function and interaction technique (IT), IT directly visible / audible, feedback 1930-1984
1984	memory, cost counter, ring volume, listening equipment, recall, calculator, clock	portable phone	push button with a beep, a repeated call can be made by pressing the recall button 1984-1990
1990	wipe out last number, muting, internal connection forward, services (e.g. tele-=shopping, tele-banking, interactive videotex), tone/pulse mode for other devices	pc, modem	making call by pressing the memory button, pc menu's with selection procedures for services
1992	hand-free call, fax, answering machine, blocking feature, hold, intercom, note pad, greenpoint system, travel information, automatic recall	telephone booth with card phone	visual feedback via display
1993	answering machines: remote controled, calls counter, memo function, Greenhopper		corrective feedback in case telephone features fail
1994	GSM system, call back system of Global Access, time indicator, hanging up system	speech output, GSM phone with information display (also for input)	info systems with speech technology and verbal choice procedures, programming numbers needs a large procedure, visual feedback about memorised number, a call can also be made by choosing last memory button that the user presses before picking up the receiver
1996	call-pin, tele-message, tele-memory, wake-up service, waiting call, call answering service, (now called: voice mail), display of incoming number, baby phone, internal call	visual attributes on display, internet, scope card	menu's choice and navigation procedures on the telephone
1997	tele-banking with creditcard	chipper for telephone cells, smartphone, creditcard, television, pc	combination of smartphone, creditcard, pc or tv needed for electronic banking

Figure 2.3: Overview of user interface changes on the telephone from 1930-1997. The years within the figure indicate the transition period between new styles.



☐ electro-mechanical style ░ display style ▒ menu style

year	conceptual operations (functionality)	interaction techniques (input/output devices)	interaction structure (procedures)
1979	video recorders for home: record and play, 8 channels, tape counter, a timer for programming	push buttons, switches	programming by pressing two push buttons simultaneously (like an audio recorder) 1979-1981
1981	8 to 26 channels, doubling, single shot, slow motion, fast motion, fast search system, go-to feature	wireless RC	RC mainly for basic functionality
1984	nullstop, super slow motion, fragment per fragment, connections (scart-plug, video camera, microphone), noise reduction, tape counter in time units, remaining tape indicator, sleep timer, insert	wheels, tools, buttons display	first type of visual feedback via display, programming via tv screen (AKAI) (type of menu) and others via display
1985	hi-fi stereo sound, one touch recording	display on RC	one press to record a program, recording has to be stopped by pressing the stop button
1988	automatic channel programming, 40-60 channels, digital video recorders		different programming procedures: step by step, automatic, reading pencil
1989	100 channels 1989-1990		
1990	programming 4 to 8 programmes at once, programming on fixed days and time, long play, fast start, blue mute, electronic clock, some RC can learn to use other devices, VPS (kind of PDC), introscan, video access search		corrective feedback: beeps when setting the start time at a later point in time than the stop time 1990-1992
1992	PDC, VPT, blank tape search	jog shuttle	programming by teletext (first form of programming via a menu structure)
1993		dip switches, little buttons on the RC	
1996	PDC and show view integrated, follow tv		automatic channel match with the tv channels via “button press” after navigation through a menu, a code and PDC activation

Figure 2.4: Overview of user interface changes on the consumer video recorder from 1979-1996. The years within the figure indicate the transition period between new styles.

with a minimum amount of functionality which made it also a typical electromechanical type of device.

Although the functionality and the interaction techniques of the television changed already in the seventies, the interaction structure remained as it was until the eighties. In the eighties the implemented programming facilities and software-based interaction techniques led to a more complex interaction structure. The one-to-one relation between a function and the corresponding interaction technique was abandoned by the display, and the interaction techniques had numerous functions depending on their mode. This new interaction style is called display style. In the eighties also the telephone and video recorder appeared to change in the same way. Note that on the telephone and video recorder all interaction style components changed in a short time, instead of gradually as seen for the television.

The moment of introduction as stated in the tables does not represent the exact moment of introduction onto the market. It reflects the first time that the feature was mentioned in the literature. Therefore, the introduction of the new style can only be expressed in terms of fuzzy periods. According to the transition periods visualized in the figures, the introduction of the display style occurred somewhere between 1978 and 1985 for the television, between 1930 and 1984 for the telephone (but more towards 1984 than towards 1930 given the evolution of most display-style features), and between 1979 and 1981 for the video recorder. We will take 1980 as the average starting-point of the display style.

Due to the influence of the computer, the interaction style changed again within 10 years. Specialized functionality focused on service, navigation, and selection procedures for the newly introduced menu structure, and visualized software-based interaction techniques dominated the user interfaces of the nineties. This interaction style will be named the menu style and was implemented on the television, video recorder, and telephone around the same time. This style is, in fact, an advanced form of the display style, as it still contains the basic elements of a display style: software based input and output devices, complex and multi-layered procedures, and a large amount of functionality.

The menu style was introduced between 1989 and 1991 for the television, between 1984 and 1990 for the telephone and between 1990 and 1992 for the video recorder. We will use 1990 as the average starting point of the menu style.

Diffusion of interaction styles

Realizing that knowledge about the interaction style changes over the years is the first step towards differentiation of technology generations based on generation-specific user interface experience in the formative years. The second step is to gather information about the diffusion time needed for each newly introduced interaction style before the majority of the people has some experience with it in the formative period. Only then the generational boundaries can be determined.

According to Baudet [1986] diffusion is the spread of innovations over users. He defines the diffusion rate as the percentage of actual users in relation to the number of potential users. In accordance, we define the diffusion rate of ICE devices as the percentage of households that possess such a device in relation to the total number of households. Normally, the diffusion rate of products as a function of time shows an S-shape [Rogers, 1983; Baudet, 1986]. Following Rogers [1983], Baudet [1986], and Sackmann and Weymann [1994], twenty percent of diffusion suggests that a majority of persons will have been confronted with such a device, either at their own home or via family, friends, work, or school.

Diffusion figures of the telephone indicate that in the Netherlands the telephone reached a diffusion of twenty percent between 1955 and 1960 [Hogesteeger, 1989]. As the telephone was the first of the three ICE devices in our research to reach the critical diffusion rate, we set the diffusion of twenty percent of the electro-mechanical style around 1955.

ICE devices such as the video recorder, television, telephone, and radio were initially developed in the electro-mechanical style, after which they appeared on the market in the display and menu style. Figures of Hansman [1996] and CBS (Central Statistical Office of the Netherlands) do not give an indication of the diffusion per new version of an ICE device. Hence, the diffusion of each interaction style cannot be directly deduced from these figures. Therefore, another indicator has to be defined for the transition from the electro-mechanical to the display style, and from the display style to the menu style.

We decided to start with diffusion figures of the CD-player, as it is one of the first devices to enter the market with a display style from the very beginning. The CD-player was introduced in 1983 and the first CD's came on the market in 1984. The CD-player encountered a diffusion of twenty percent in about five years in the Netherlands. In 1988, twenty-five percent of the households in the Netherlands owned a CD-player [Hansman, 1996]. The display style was introduced on televisions and video recorders somewhat before 1983. The five years of diffusion found for the CD-player has also been taken for the display style (1980 + 5 years) and the menu style (1990 + 5 years). Hence, it is expected that the majority of people will have experienced the display style at their own home or elsewhere from 1985 on and the menu style from 1995 onwards.

Formative period

The formative years are operationalized in this study as the period before the age of 25. The technology generations, based on user interface experience in their formative period can then be distinguished by calculating which cohorts did not

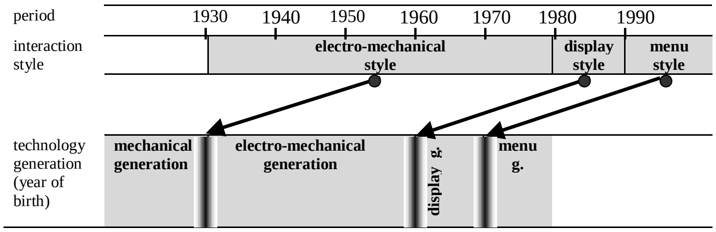


Figure 2.5: Overview of interaction styles and technology generations. The dot at the tail-end of the arrow depicts at which point in time it is hypothesized that the diffusion of an interaction style has reached the 20 percent diffusion rate. The arrow indicates when a new technology generation starts. The vertical band between two technology generations indicates the fuzzy generational boundary.

experience any of these interaction styles before the age of 25, and which cohorts experienced the electro-mechanical, display, or menu style in that period. Thus, twenty-five years have to be subtracted from the year of each interaction style at which 20 percent of diffusion was reached. The outcome defines the first birth cohort that belongs to a new technology generation.

Technology generations

Based on these estimates we could distinguish three different technology generations: the electro-mechanical generation (born 1930-1960), the display generation (born 1960-1970) and the menu generation (born after 1970). According to our historical analyses, the age group born before 1930 was not acquainted in their formative years with electrical ICE devices such as the telephone, video recorder and television. Therefore this group is part of a separate generation, which is defined here as the mechanical generation (born before 1930). Figure 2.5 shows the hypothesized relationship between the interaction styles, their period of implementation and the corresponding technology generations.

Note that the menu style was introduced shortly after the display style (10 years interval). This shortens the pure display generation considerably. Consequently, it is questionable whether the display generation can be empirically differentiated from the menu generation. Later on in chapter 3 and 4 we will merge these two generations (display and menu) to one generation, called the software generation.

2.2.5 *Conclusions*

Three ICE interaction styles were introduced in the second half of the 20th century: (1) the electro-mechanical style, available between 1930 and 1980, (2) the display style, available from 1980 to 1990, and (3) the menu style, implemented from 1990 until now.

Although the historical analysis has indicated which interaction styles came onto the market, the precise moment could only be estimated, and the diffusion time of each interaction style could not be derived from the historical analyses. Literature indicates that the electro-mechanical style reached a diffusion of 20 percent around 1955 [Hogesteeger, 1989]. Diffusion figures of the CD-player suggest a diffusion period of 5 years for the display and menu style [Hansman, 1996]. These diffusion figures should be verified to distinguish technology generations that are based on interaction style experience acquired in the formative years. The next section will focus on this problem.

2.3 Technology generation boundaries

2.3.1 *Introduction*

The technology generation boundaries, as depicted in figure 2.5 should be verified empirically to analyse the accuracy of the boundaries. For this, a survey has been executed to measure the diffusion of each ICE interaction style in the subjects' formative years. These figures were then related to the technology generation boundaries as based on historical data analyses and diffusion figures found in the literature. The extent to which the figures match reflects the accuracy of the predefined boundaries.

The survey was only executed for subjects beyond their formative years. Consequently no persons from the menu generation have been interviewed. Hence, only the boundaries between the mechanical, electro-mechanical, and display generations can be verified here.

2.3.2 *Method*

Subjects

1015 Subjects were interviewed in 1998 in the Netherlands. The age range was between 30 and 79, covering the display, electro-mechanical, and mechanical generations. Hence, the menu generation will not be covered in this study. The age distribution, and the balance between the female-male population and educational levels per age group (see figure 2.6 and table 2.1) resembled the normal Dutch population.

Material and procedure

The survey was executed in cooperation with A.G. van de Goor and H. Becker of Utrecht University. Our part concerned a structured face-to-face interview with multiple-choice questions concerning the availability of a list of ICE appliances at the subject's parental home and experience of each appliance before the age

Table 2.1: The number of subjects per educational level, depicted per age cohort range.

educational level	30-44 years	45-64 years	65-79 years
master degree	14	17	9
bachelor degree	43	49	16
pre-university education	40	16	10
intermediate vocational education	79	83	21
intermediate secondary education	48	66	58
lower vocational education	87	114	63
primary school	14	62	89
primary school not finished	1	4	12

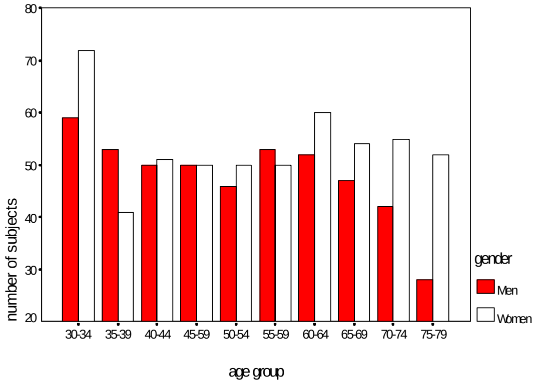


Figure 2.6: Age distribution of the subjects interviewed in this study, depicted per gender.

of 25. The entertainment appliances concerned the radio, black and white television, colour television, remote control, teletext, tape recorder, cassette recorder, video recorder, and compact disc player. Information-communication appliances concerned the wired and wireless telephone, mobile phone, fax, answering machine, personal computer, CD-rom, Internet, and electronic mail. The subjects were interviewed at home.

Definition of technology generation boundary

The generation boundary is defined as the point in time at which birth cohorts minimally showed 20 percent of diffusion of a certain interaction style in their formative period. According to our historical data analyses, the telephone was one of the first devices implemented with an electro-mechanical style. Consequent-

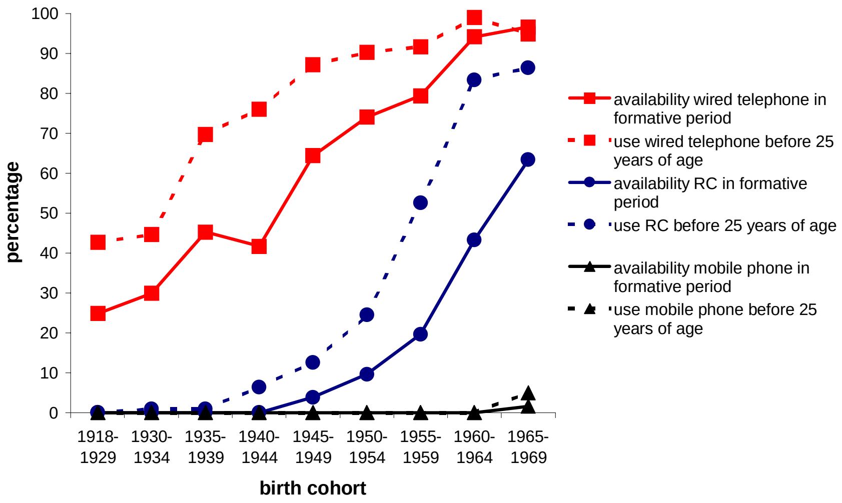


Figure 2.7: Diffusion figures per birth cohort of the wireless remote control (RC) and the wired telephone, based on availability in the parental home and experience in the formative period (before the age of 25).

ly, for the boundary between the mechanical and electro-mechanical style, the diffusion figures will be used of the telephone at the parental home in the subject's formative period. As discussed earlier in this chapter, the CD-player has been used at first to calculate the diffusion time of the display style. In this survey we were able to measure also the earliest feature of the display style, namely the wireless remote control. According to the historical analyses the wireless remote control came onto the market in 1974 for the television and in 1981 for the video recorder. Therefore, the boundary between the electro-mechanical and display generation will be analysed by the diffusion figures of the wireless remote control at the parental home in the subject's formative period.

In this study, the boundary between the display and menu generation can roughly be checked by the availability of the mobile phone in the formative years.

2.3.3 *Results*

Boundary between the mechanical and electro-mechanical generation

The boundary between the mechanical and electro-mechanical generation was estimated to be located around the birth year of 1930. According to the survey results the availability of the wired telephone at their parental home already showed a diffusion of 25 percent for subjects born before 1930, and even 40 percent of this group mentioned to have used it before the age of 25 (see figure 2.7).

Boundary between the electro-mechanical and display generation

The 20 percent diffusion of the display style was checked by the diffusion of the remote control. The purchase diffusion of 20 percent of the wireless remote control was reached at the birth cohort of 1960 (see figure 2.7). The percentage of subjects born in 1960 and later that experienced the display style was even higher. Eighty three percent of these subjects mentioned to have used it before the age of 25 (see figure 2.7).

Boundary between display and menu generation

According to the theory, the population interviewed in this survey should not be part of the menu generation. As a rough check, mobile phone use and its availability in the formative years is measured and depicted in figure 2.7. As can be seen very few people born before 1970 were confronted with it in their formative years. Unfortunately no detailed questions were asked about the availability and use of the personal computer in the formative period. The availability of the Windows version on the PC could have been even a more accurate measure for the boundary check.

2.3.4 *Discussion*

Boundary between mechanical and electro-mechanical generation

The results partially confirm the defined generations via historical data. A diffusion of more than 20 percent is seen for the birth cohorts between 1918 and 1929 concerning the availability of the wired telephone (25 percent) at their parental home. This does not correspond with lower telephone diffusion figures mentioned in the literature [Hogesteeger, 1989]. This mismatch may be due to the fact that Hogesteeger calculated the diffusion of products on a certain moment in time for the total population, whereas in our study the diffusion figures were calculated per age group. Furthermore, we analysed the diffusion of a certain device within the subject's formative years. The figures of Hogesteeger are not limited to a certain life phase. Moreover, the figures in our study referred always to families with children. The study of Van de Goor and Becker [2001] indicated that families with children have more consumer products at home than households without children.

The availability outcomes suggest that the boundary between the mechanical and electro-mechanical generation should be placed earlier in time than 1930 as found by the historical analysis. On the other hand, the data about the subject's experience in the formative years indicate that the telephone was not used by the majority of the subjects belonging to the birth cohorts born before 1929. This latter aspect suggests that the predefined boundary between the mechanical and electro-mechanical generations, based on user interface experience of the majority of people, should be placed even later in time: around the birth cohort of 1935. Moreover, another factor should be kept in mind. In our historical data analyses the diffusion of radio was not included. Empirical data show that the diffusion of the radio [Baudet, 1986] at the subject's parental home played a major role for birth cohorts born before 1929 concerning their knowledge about electrical ICE user interfaces in the formative period. Therefore, such data have to be included as well. On the basis of this argument the boundary between the mechanical and electro-mechanical generation is expected to be found before 1930 after all. Which birth cohort corresponds with the boundary as suggested here cannot be checked as subjects older than 79 were not interviewed. In summary, all these arguments indicate that the boundary between the mechanical and electro-mechanical generation cannot be set unambiguously around 1930. Further research is needed to analyse the appropriate one.

Boundary between electro-mechanical and display generation

The boundary between the electro-mechanical and the display generation that can be derived from the survey outcomes resembles the predicted boundary around the birth cohort of 1960. Note that that estimate was based on historical data that described a fuzzy period in which the display style was introduced to the market and the calculation of the 20 percent of diffusion via diffusion figures of the CD-player [Hansman, 1996]. Despite this rough estimation it resembles the much more precise measurement of the boundary between the electro-mechanical and display generation that was operationalized via diffusion figures about the first display style feature, namely the wireless remote control. Following Rogers' rule of thumb 20 percent of diffusion indicates that at such a point in time the majority of the population will be acquainted with the product [Rogers, 1983]. This is indeed true for the wireless remote control and thus for the display style: a majority (83 percent) of subjects born in 1960 and later mention to have used it in their formative period.

Does this result mean that a diffusion time of five years is a good estimate for the critical diffusion rate of the display style in general? If most people have experienced the wireless remote control around 1985 as the survey results suggest, it would mean that the RC of the television has been diffused very slowly (from 1974 to 1985), whereas the RC of the video recorder diffused within the estimated five years (from 1981 to 1985). This is not so surprising as the colour television was already at home in most households when television sets with RC came on the market. The average life time of a television is about 12 years. This means that people would massively have to exchange their working television for a new one with RC. This was of course not the case. This situation was not the case for the video recorder. Most people bought their first video recorder in the eighties. Video recorders with RC were then probably favoured over the older models. The period between 1981 and 1985 was also the starting period in which most consumer products were implemented in this so-called display style. Therefore, the period between 1980 and 1985 is indeed expected to be the core period of diffusion of the display style.

Boundary between display and menu generation

As the subjects interviewed in the survey were 30 years or older, the boundary between the display and menu generation could not be verified at all. The diffusion results of the mobile phone, which is one of the first devices that came onto the market directly in the menu style, suggest that the boundary between the display and menu generation should be placed somewhat later in time. This outcome is however not conclusive as other ICE products, such as the personal computer and the television, may have been implemented somewhat earlier in the menu style than the development of the mobile phone. However, their appearance in the person's formative period has not been checked in detail.

2.3.5 *Conclusion*

From the four predefined technology generations (mechanical generation, electro-mechanical generation, display generation and menu generation), two generations have been identified via the survey. The electro-mechanical generation overlaps the predefined mechanical generation, and therefore covers at least the birth cohorts between 1918 and 1959. The display generation covers the birth cohorts of 1960 and later. Due to restrictions of the survey (only subjects born before 1969 were interviewed), it could not be tested whether and which birth cohorts belong to the menu generation.

2.4 General conclusions

Based on historical analyses at least three interaction styles have been implemented in the second half of the $20^{\rm th}$ century: the electro-mechanical style (1930-1980), the display style (1980-1990), and the menu style (1990 and later). Based on these user interface categorizations and survey analyses, at least two technology generations can be distinguished: the electro-mechanical generation (born before 1960), and the display generation (born in 1960 or later).

3

Using layered user interfaces: Age and technology generation effects

In two experiments it is investigated in this chapter why older adults encounter more difficulties using current devices than younger adults. It is hypothesized that two factors play a role: (1) age-related declines in cognitive abilities, and (2) generation-related lack of experience with present (multi-layered) user interfaces. In the first experiment, the performance of four age groups (25-35, 40-50, 50-60, and 65-75) was compared using a simulation of a single-layered and a two-layered videophone interface. Age group differences observed in error performance indicated a generation effect. Task duration analyses only showed an age effect. In the second experiment, a similar design was used. The focus was now on a simulation of a three-layered mobile phone interface. Again, a technology generation effect was found for error performance and an age effect for task duration, which suggests that these effects persist in other, more complex tasks and devices.

3.1 Introduction

Many adults seem to have difficulties using present-day consumer products with their extensive functionalities and many options [Stewart, 1992; Freudenthal, A., 1998]. At least three factors can be discerned that may potentially contribute to these difficulties: complexity of the user interface, age-related changes in cognitive abilities, and generation-related differences in experience with technology.

The objective of the present study is to investigate to what extent age-related cognitive changes and generation-related past experience with technology, together with the user interface of the consumer product, affect user performance. To this end, the behaviour of four age groups (25-35, 40-50, 50-60, and 65-75 years of age) has been analysed by means of two experiments. In the first experiment, two simulations of a videophone interface were presented, one with a single-layered and one with a two-layered structure. In the second experiment, a simulation of a three-layered interface of a mobile phone was presented.

Section 3.2 provides the theoretical framework of the present study, leading towards the discussion on how age and generation effects can be experimentally separated. Sections 3.3 and 3.4 describe and discuss the two experiments. This chapter finishes with a general discussion and some implications for design and further research.

3.2 Theoretical framework

3.2.1 *Impact of user interface complexity*

Complexity of the user interface could be one of the reasons why current consumer products are difficult to use. In order to describe the complexity of a user interface, it is useful to have a specification of the user interface, in particular of the associated *interaction style*. This emphasis on interaction style is supported by research on human-computer interaction showing the influence of interaction style on user performance for software user interfaces (e.g. Rauterberg [1992] and Wiedenbeck and Davis [1997]). According to Colbert [1997] an interaction style comprises a specification of re-usable application controls, a specific selection and design of these controls, and a number of general design principles. This definition has been extended with the concept of directness of manipulation, where direct and indirect manipulation styles serve as extremes on a continuum (see for example Schneiderman [1983], Wiedenbeck and Davis [1997], Colbert [1997], and Hutchins, Hollan and Norman [1985]).

Recently, De Vet and De Ruyter [1996] described interaction styles of audio products in a way that offers a structured approach to differences in styles and their impact on user performance. They proposed that an interaction style supports generic user-task options (*conceptual operations*), organizes such options according to a specific operational procedure (*interaction structure*), and uses specific interaction procedures for input / output devices (*interaction techniques*). For example: a mouse-pointing device (= interaction technique) is used to select a file (= a conceptual operation) by clicking several icons (= interaction techniques) via a hierarchical menu composed of a fixed dialogue flow (= interaction structure). User interfaces can always be described in these terms and, accordingly, the impact of each aspect on the user's behaviour can be analysed by manipulating them. We, therefore, adopted this way of identifying the complexity of user interfaces. In particular, the present study focuses on the impact of the interaction structure. We assume it plays a key role in the user's understanding of the user interface organization, and thus on the acquisition of procedural knowledge needed to use a product without difficulties.

Docampo Rama and Van der Kaaden [2001] analysed how the interaction style of consumer products has changed over the past century. In the beginning of the eighties one major change has been identified: the style changed from an electro-mechanical style to a software style. This change is characterized by the interaction techniques evolving from mechanical attributes (e.g. push buttons, switches and rotary dials) to software-based ones like displays, touch buttons, and wireless remote controls. Furthermore, the functionality available changed from basic functionality to many specialized functions, whereas the interaction structure changed from single-layered to multi-layered. This means that the former was organized in breadth, such that all functionality was directly visible to the user, whereas the latter style is hierarchically organized, and temporarily less relevant functionality is often hidden. In addition, most input and output devices of multi-layered consumer products have many functions depending on their mode, and have been implemented without providing direct feedback about their state.

In this study the complexity of software style devices will only be analysed as far as it concerns its layered interaction structure. A layered control panel often fails to give information about the functions of each control and the status of the information behind it, nor does it give an overview of the available functionality. Consequently, such user interfaces may impose a large demand on the user's reasoning ability for figuring out new procedures. Furthermore, it requires a good working memory [Baddeley, 1986] to keep relevant user interface information available for the execution of the goal in mind. Especially the *visuo-spatial sketch-pad* [Logie, 1995] is expected to be heavily involved when using multi-layered interfaces, as no cues are given by the control panel about the spatial location of each function. User interfaces organized in breadth are expected to rely much less on these cognitive abilities.

This latter suggestion is supported by a number of studies. Gomez, Egan, Wheeler, Sharma and Gruchacz [1983], for example, found that visuo-spatial ability correlated strongly with the performance using a screen-based and a text-based software editor. Similarly, Detweiler, Hess and Ellis [1996] compared the user's accuracy using a window display (which has a multi-layered structure) and a grid display (which resembles a single-layered structure). A window display appeared to be more difficult to use than a grid display due to its lack of information about the spatial location of functions. Beside that, users were less accurate on a window display relative to that of a grid display when the number of values they had to keep track of increased. Hence, it seems that a window display or multi-layered interface indeed induces a greater working memory load, and is therefore more difficult to use than a grid display or single-layered interface. Landauer and Nachbar [1985] showed that the user's mean response time for decision and movement is shorter when user interfaces are structured in breadth than when structured in depth.

3.2.2 *Impact of age*

There is substantial evidence (e.g. Craik and Bosman [1992]) that some cognitive abilities for learning complex systems change with age. This certainly holds for the user's reasoning ability and working memory for spatial information. Welford [1980] and Hultsch and Dixon [1990] state that it is the ability to hold information in working memory that tends to decline with age. Salthouse and Skovronek [1992] and Howard and Howard [1997] explain these working memory restrictions in terms of an age-related reduction of processing capacity. See Salthouse [1991] for an overview of changes in spatial abilities with age. Most of these abilities show an almost linear decline with age starting around the age of thirty to forty. Normative studies of the Raven Matrices test [Raven, 1960] indicate that abstract reasoning with visual material declines almost linearly starting from the age of thirty onwards [Salthouse, 1991]. Longitudinal studies showed that declines in the ability of reasoning and spatial orientation, as measured by parts of the WAIS, are found in people only in their late sixties [Schaie, 1983]. Hence, the trends that are found in ageing literature about cognitive changes vary between an almost linear decline from the age of 30 onwards to a monotonic decline that can be characterized by a shallow slope until the age of 60 and a much larger decline after that age.

Studies on complex systems and ageing have shown that cognitive changes with age indeed may explain observed age differences. Salthouse, Hambrick, Lukas and Dell [1996], for example, found that the age differences in performance involving complex parallel activities had a 70 percent overlap with age-related differences in processing speed. Freudenthal, T.D. [1998] has shown that there also are age differences on the level of reasoning, as defined by the GEMS model of Reason [1990], when exploring a new interactive system. With increasing age the performance on the highest level of reasoning decreases and the performance on a trial-and-error basis increases. Furthermore, he showed that age-related changes in spatial ability explain part of the latency differences found between older and younger subjects working with a menu structure. Similarly, Czaja and Sharit [1998] showed that visuo-motor skills and memory were good predictors of the age-related differences in quantity and quality of work done during a data entry task. In this study it is therefore expected that age differences will be seen when working with complex devices.

According to the age-complexity hypothesis (e.g. Cerella, Poon and Williams [1980]), performance differences between younger and older adults increase with task complexity. Results of Gomez, Egan, Wheeler, Sharma and Gruchacz [1983] already confirmed this tendency using screen-based versus text-based editors. Consequently, the largest performance difference between younger and older adults is expected using a multi-layered interface and a less pronounced age difference when using a single-layered one.

3.2.3 *Impact of generation*

Next to general cognitive abilities, also specific experience is expected to play a role using a complex system. The user needs to acquire declarative and procedural knowledge [Anderson, 1983] about the functionality, structure, and controls of the system. The greater the knowledge base of a person on a topic, the easier it is to encode, understand, integrate, and remember new relevant information [Searleman and Douglas, 1994]. As all this information is stored in long-term memory, the user is able to retrieve and use this information whenever he/she uses another device. One advantage of such a mechanism is that learned features and procedures can be reused in the new situation without effort.

One important question is to what extent knowledge can be generalized to a new situation. Studies show that prior experience with a device and its user interface has a positive influence using a new device with a similar user interface [Kelley, 1996]. Transfer of knowledge can also be negative [Singley and Anderson, 1985] when users are compelled to abandon their habits and restart learning new procedures.

People experience many different types of technology during their life time. Not all the knowledge acquired is expected to be that useful, as the user interface has changed a lot over the years. Consequently, only relevant knowledge about complex multi-layered user interfaces is expected to facilitate the use of current devices. Czaja and Sharit [1993] already showed that age differences in the usage of interactive systems are related to lack of experience of older users with present-day devices. Similarly, Freudenthal, T.D. [1998] showed that age differences were most pronounced when young and old users were not informed about the possibilities and procedures of the interactive system to be used.

Social studies concerning cohort or generation analyses [Glenn, 1974; Cutler and Kaufman, 1975] have indicated that persons after young adulthood are less likely to become susceptible to change in attitudes, norms and values, and behaviour. The period of young adulthood and adolescence, operationalized as the period between 10 and 25 years is also known as the *formative period* and plays a key role in the formation of these attitudes and behaviours. In this period, people have to master a number of crucial transitions, for example from school to work, and from parental home to living independently. During these events, people acquire important values, norms, and technical skills that influence the way they behave during the rest of their life [Becker, 1992]. Reasons often given for the attitudinal and behavioural rigidity after the formative period are, among others, neurological change, decreased social and geographic mobility, and consequent changes in social contacts.

According to Sackmann and Weymann [1994] and Becker and Van de Goor [1997] also technology experience acquired in the user's formative period has a great impact on the user's way of handling current technology. Sackmann and Weymann [1994] have found that people, who experienced the availability of the same types of consumer products during their formative period, show similar technology usage behaviour today. Such a group has been identified as a *technology generation*. In their theory, different technology generations behave differently with respect to technology, due to differences in technology experience during their formative period.

Generation-specific technology experience could also induce differences in using current consumer products. Again the formative period is expected to play a key role here. Child development studies indicate that a person's cognitive abilities reach their highest level, such as hypothetico-deductive reasoning, during adolescence (Piaget and Inhelder [1973], Keating [1980], and Neimark [1982]). In this period the person's social relationships and other life experiences both influence and are influenced by cognitive advances [Sroufe and Cooper, 1988]. Consequently, the formative period is the optimal phase of life in which complex systems and their interaction style can be understood and learned to use in all its facets. Therefore, it is hypothesized that lack of knowledge about current user interfaces as well as habits acquired during the formative period may complicate the older user's learning process in the new situation. Birth cohorts that have only experienced electro-mechanical user interfaces are expected to encounter more difficulties using current software style devices than cohorts that experienced the introduction of present-day user interfaces during their formative period.

The problem then is to determine which birth cohorts only experienced electro-mechanical interfaces, and which experienced software style products. Following Rogers [1983], Baudet [1986], and Sackmann and Weymann [1994], we operationalize the beginning of the period in which these current user interfaces were widely introduced as the moment at which more than twenty percent of the population possessed a device with such a user interface. At this point of diffusion, it is likely that persons who do not possess such a device will have been confronted with it via relations, friends, at work, or at school.

Video recorders, televisions, telephones, and radios were initially developed with an electro-mechanical user interface. In the eighties, consumer products appeared on the market in a software style, and could often be remotely controlled. However, in the standard literature it is not known how many and which people bought the software style version. Thus, how do we define a suitable indicator for the transition?

Docampo Rama and Van der Kaaden [2001] carried out a survey on 1015 Dutch subjects to analyse the time course of diffusion of current user interfaces. Questions were asked about the availability of software style consumer products, such as the remote control and teletext[1], at their parental home and their experience with it before the age of 25. The outcome suggested that in 1985 more than 20 percent of the subjects had software style devices at their parental home and more than 80 percent of the subjects had used them before the age of 25. Therefore, birth cohorts born in 1960 or later will be referred to in this study as the *software generation* (abbreviated as *S-generation*). This means that birth cohorts born before 1960, did not experience the presence and availability of software style interfaces in their formative period. In this study, this latter group will be called the *electro-mechanical generation* (abbreviated as *EM-generation*). It is assumed that the latter generation encounters more difficulties using a software style interface than the S-generation.

3.2.4 *How to distinguish an age effect from a generation effect*

On the basis of the above it is hypothesized that the learning performance of a person using a new device, with a given complex user interface, is a function of (1) age, referring to the user's cognitive abilities that are susceptible to changes with age, such as visuo-spatial working memory and reasoning capacities, and (2) technology generation, referring to the user's skills, acquired through generation specific technology experience in the formative years. Note that this does

---

[1] Teletext is also one of the first software style features on Dutch television and started in 1978. It is a textual information system.

not mean that other factors are not involved. One can think of cognitive abilities such as word fluency, verbal meaning [Schaie, 1985], and vocabulary [Craik and Bosman, 1992], technology experience gained after the formative period, and other individual differences. Some of these factors are assumed to be less sensitive to changes with age while others are not of major importance to the present study.

Empirically, it may be difficult to dissociate a generation effect from an age effect, because both age and technology generation are measured by the person's year of birth. According to standard sociological methods, age effects can be distinguished from generation effects by analysing the performance across age cohorts; see Glenn [1977] for an extensive explanation on this topic. The effect is expected to be age-related if the data show a continuous change of performance with age. A discontinuous change suggests a cohort, or generation effect. This type of analysis requires a large number of subjects with a dense coverage of a wide age range. In general, such a design is not realistic in human-system interaction studies.

In the present study the problem of separating age and generation is tackled by using four age groups. One group is between 25 and 35 years old and part of the S-generation. The other three age groups (respectively 40-50, 50-60, 65-75) belong to the EM-generation. As the complexity of the interaction structure is expected to influence the performance of everyone, these groups are instructed to work with a software style user interface featuring a single-layered and a two-layered interaction structure (experiment 1) or a software style user interface with a three-layered interaction structure (experiment 2).

The effects of age and generation are to be separated by analysing the performance across age groups. If only an age effect is involved, an age-related decline in performance is found, operationalized as an increase of errors and/or task duration. Such age decline is expected to be more prominent for the multi-layered interface than for the single-layered interface. We hypothesize that the performance declines either (1) linearly with age, starting from the age of thirty, as was indicated by cross-sectional studies mentioned in the theoretical framework of this study, or (2) monotonically with age, perhaps with a flat slope between the age of 30 and 60 and a sharp decline after the age of 60, as shown by some longitudinal studies [Salthouse, 1991]. The left-hand panel of figure 3.1 depicts these hypothesized relations. If only a generation effect is involved a discontinuous change between two plateaus is expected at about the age of 40. Note that this tendency does not correspond at all with cognitive ageing trends mentioned earlier in this chapter. The right-hand panel of figure 3.1 shows the generation effect. A third option is that both age and generation effects are involved. Figure 3.2 depicts

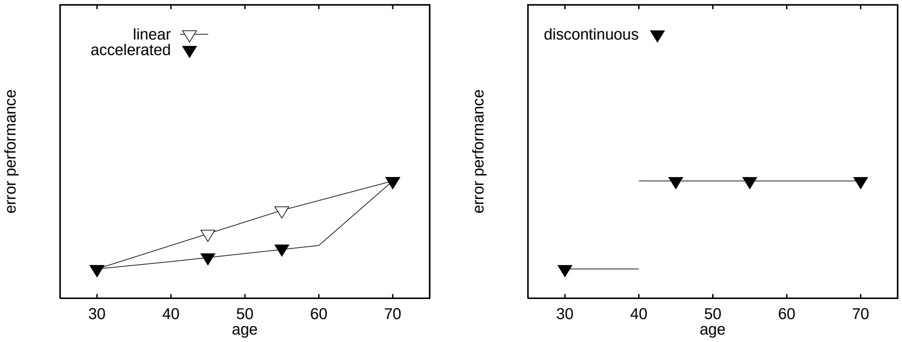


Figure 3.1: Left-hand panel: The hypothetical linear or accelerated performance trends across age groups, if there is only an age effect. Right-hand panel: The hypothetical plateaus, referring to a discontinuous change in performance, if there only is the technology generation effect as explained in the text. The symbols denote the schematic outcomes of the present study under the assumption of either an age or generation effect.

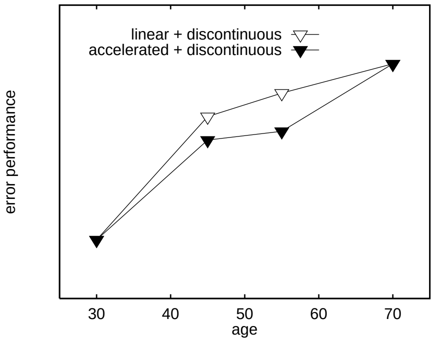


Figure 3.2: The hypothetical trends across age groups, if both an age and a technology generation effect play a role in influencing performance. It is assumed that the age and technology generation effects are additive factors. The symbols denote the schematic outcomes of the present study under the assumption of an age and a generation effect.

how performance is affected if age and generation are additive factors.

3.2.5 *Scope of the present study*

Two experiments were conducted. The first one examined whether: (1) the performance indicates a continuous or discontinuous performance decrease with age, or a combination of both, and (2) users learn to use a single-layered interface faster than a two-layered one. A second experiment focused on the generalizability of the effect found by using a three-layered user interface. The decrease in performance was measured in terms of an increase in error performance and task duration. Hence, error performance and task duration were the dependent variables.

3.3 Experiment 1: one and two-layered interfaces

The hypotheses in the first experiment were investigated by running software-based simulations of a videophone provided with a single-layered or a two-layered interaction structure. This implies that the conceptual operations and interaction techniques are kept constant, but the interaction structure is manipulated as far as it concerns its layers.

The task of the subjects was to contact a person with a prescribed name in 3 conditions, manipulating the presence of video and subtitling. The level of performance was measured by errors and task duration. The error performance was measured by the number of *mode errors*, where a mode error is defined as an action that does not have any effect on the device as the action is not executed in the correct state (for example, trying to make a call while the telephone is switched off). Inexperienced users in particular make these errors due to their erroneous expectations about the *conceptual model* of the system, i.e., the way it works [Johnson, 1990]. In contrast to certain other types of errors, a mode error confronts users directly with the failing action. Consequently, the occurrence of mode errors serves as a good indicator for the user's performance problems.

Subjects were expected to have more difficulties forming a conceptual model of the system, and thus making more mode errors and operating more slowly, with a two-layered than with a single-layered interface. Learning may be expected by repetition and also by knowledge transfer from one user interface to another. Therefore we chose for two trials with one user interface, followed by two trials with the other user interface. Relatively fewer mode errors were expected for interaction with a second user interface.

3.3.1 *Method*

Subjects

Eighty subjects took part in the experiment. The age range was 25 to 75 years. The subjects were assigned to four age groups. The age of the youngest group ranged from 25 to 35 (representing the S-generation), with a mean age of 30. The age of the first middle group ranged from 40 to 50, with a mean age of 45; the second middle group ranged from 50 to 60, with a mean age of 55. The age of the oldest group ranged from 65 to 75, with a mean age of 68. The latter three age groups represent the EM-generation.

None of the subjects had ever used a videophone before. Educational level was confined to no more than a high school diploma, resulting in a homogeneous experimental population in this respect. An equal number of female and male subjects was selected per age group to control for possible gender differences. Finally, the subjects did not have any self-reported problems with manual movements, hearing, or vision. In addition, the latter aspects were checked by the experimenter just before the beginning of the experiment by asking questions about the visibility of the videophone interface and the recognizability of the sounds.

3.3.2 *Design*

The independent variables were the user interface (single-layered or two-layered), the trial per user interface (first or second), the user interface order (interacting first with the single-layered interface and after that with the two-layered interface or vice versa) and the age group (25-35, 40-50, 50-60, or 65-75). The experimental design was a 2 x 2 x 2 x 4 counterbalanced block design. The user interface order and age served as between-subjects variables. Per age group, 5 male and 5 female were randomly assigned to one order, and the others to the other order. The user interface and the trial per user interface were the within-subject variables. The dependent variables were the total number of mode errors and total task duration per trial.

Materials

Two software-based videophone user interfaces were simulated using Microsoft Visual Basic 4.0 in the Windows 3.11 environment. Figures 3.3 and 3.4 depict the simulated single-layered and the two-layered videophone interface respectively. Three tasks were implemented on both videophones. These tasks represented part of the functionality of the simulated videophone.

Both user interfaces had three switches. One switch turned the device on or off, one selected the channel (audio, or audio and video), and the third switch activated subtitling. Furthermore, three buttons were added that functioned as memory buttons for telephone numbers. These buttons and switches could only be used in a certain order. This order was exactly the same for both types of user interface. If the subjects confused the order in which these buttons and switches had to be manipulated this resulted in a mode error. On the first layer of the two-layered interface (see figure 3.4) three rectangular buttons A,B,C can be seen. These buttons corresponded with the spatial position of the three switches as presented on the single-layered interface. If one of these buttons was pressed, only the corresponding switch popped up on a display, placed in the middle of the device. The rectangular buttons did not have to be pressed in a fixed order. Hence, visualization in itself of a switch via such a button did not result in a mode error. Both types of user interfaces used the same type of switches, buttons, con-

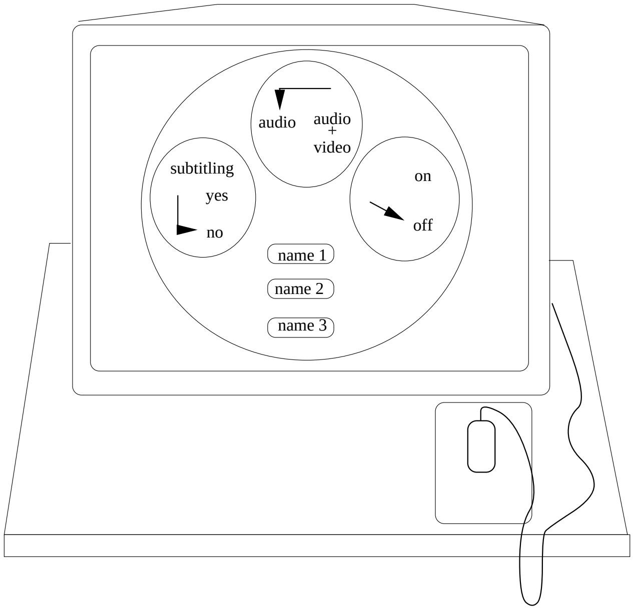


Figure 3.3: Experiment 1: single-layered videophone interface. Words were presented in Dutch.

trasting colours, fonts, font sizes, and words (except for the names of people on the memory buttons). All buttons, switches, and words were clearly and equally visible. Both videophone interfaces had to be controlled by a mouse-pointing device. Subjects could click on buttons and switches to activate them. The experiment itself was run on a laptop Pentium 130. The subjects were seated in a quiet room that was well lit without specular light reflection on the laptop screen.

Procedure

The experiment consisted of four trials. Each subject exercised with one user interface during the first two trials, and worked with the other user interface during the last two trials. Per age group half of the subjects started with the single-layered interface and ended with the two-layered interface and the other group worked the other way around. There were no time breaks between the trials. Each trial consisted of carrying out three videophone tasks calling a certain person: (1) without enabling the video screen, (2) after enabling the video screen, and (3) after enabling the video screen and having activated the subtitling of incoming speech. These tasks were always executed in this order. Nine steps were needed on the single-layered videophone to carry out these tasks, whereas fifteen steps

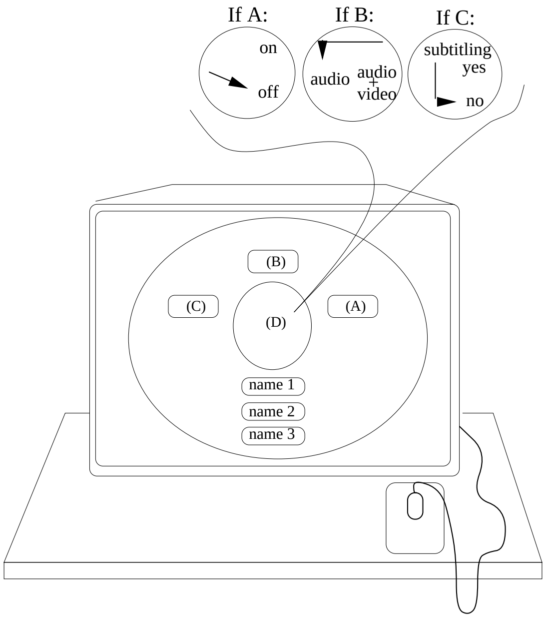


Figure 3.4: Experiment 1: two-layered videophone interface presented in experiment 1. Words were presented in Dutch. The upper part of the figure depicts the switch that appears on display D when button A, B, or C was pressed.

were necessary on the two-layered interface. Before starting the first trial, all subjects received a general introduction and practiced with the mouse-pointing device by means of two exercises. Specific instructions were given before each user interface was employed. After the final trial of the experiment, each subject was asked whether he/she had prior computer and mouse experience.

3.3.3 *Results*

The optimal performance during the four trials is 48 steps. On average 87 steps[2] were needed by the youngest group, 164 steps by the age group of 40-50, 104 by the group of 50-60 years and 129 by the oldest group. On average the youngest group finished the four trials in 5 minutes and 29 seconds, the group of 40-50 years in 10 minutes and 37 seconds, the group of 50-60 years in 12 minutes and 38 seconds, and the oldest group in 17 minutes and 5 seconds.

---

[2] The total number of steps needed corresponds with the sum of the minimal number of actions, the total number of mode errors and the total number of redundant steps needed by the subjects.

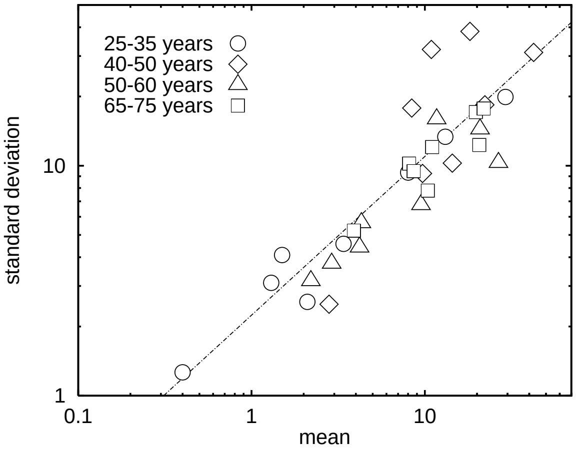


Figure 3.5: Experiment 1: Standard deviations in function of the average number of mode errors. Each point reflects the average of twenty subjects. Data indicate the same trend for the four age groups. Most highest scores within each age group correspond to the first attempts, whereas most lowest scores correspond to subsequent attempts.

Mode error transformation

Each subject took part in 4 trials. Combined with the 8 combinations of age and order, this resulted in 32 conditions (8 conditions per age group) with 10 scores of mode errors per condition. The mean and the standard deviation calculated for these conditions show that the standard deviation increased with the average number of mode errors (figure 3.5). The overall relation can be described by the following function:

$$\log s(x) = 0.35 + 0.69 \log \bar{x} \quad (r = .89)$$

Hence, the standard deviation ($s(x)$) is a power function of the mean ($\bar{x}$) with the power of 0.69. For further analyses like ANOVA repeated measures, the standard deviation should be independent of the mean number of mode errors. This can be accomplished by raising the raw data to the power of .31, i.e. 1 - .69; see Winer, Brown and Michels [1991] for a detailed explanation. Accordingly, for further analyses all raw scores were transformed by raising them to the power of .31. From here on only the monotonically transformed data will be used. The analysis of the untransformed data showed similar results [Docampo Rama, 1997].

Mode error analyses

A 2 x 2 x 2 x 4 ANOVA mixed repeated measures (user interface by trial per user interface by order by age) was carried out with the transformed number of mode errors per trial as the dependent measure. Mauchly's test of Sphericity indicated that the variance of the groups indeed was homogeneous (Mauchly's W= 1.0). In the ANOVA mixed repeated measures the main effects of age (F(3,72) =4.31, $p = .008$, power= .85), user interface (F(1,72) =30.8, $p = .000$, power= 1.0), trial-per-user interface (F(1,72) =122.88, $p = .000$, power= 1.0), and order (F(1,72) =7.82, $p = .007$, power= .80) were significant. Moreover, the interactions of order by user interface (F(1,72) =94.38, $p = .000$, power= 1.0), order by trial-per-user interface (F(1,72) =4.38, $p = .040$, power= .46) and the four-way interaction (F(3,72) =2.77, $p = .048$, power= .28) were found to be significant as well. The following sections explain the results of the analysis on the transformed data in detail.

Age and generation

Figure 3.6 depicts the transformed error performance per user interface and age group averaged across trial and order. The four age groups are represented by their mean age. The figure shows that (1) the two-layered interface elicited more errors than the single-layered interface and (2) the youngest age group makes fewer mode errors than the other three age groups. The three oldest groups did not show an age-related increase of transformed mode errors. This pattern suggests a discontinuity between the age of 30 and 45 as mentioned in section 3.2.4, which can be referred to as a generation effect. The Helmert Contrast analyses confirmed this effect: only the age group of 25-35 differed significantly from the three older age groups (t= -3.47, $p = .00$). The three age groups from the age of 45 on showed a similar error performance (t= 0.88, $p = .38$ for the group of 40-50 years versus the older groups and t= 0.58, $p = .56$ for the difference between the two oldest groups). Figure 3.6 also suggests that there was no difference in the relation between age and transformed error performance for the single and two-layered interface. This was confirmed by the ANOVA repeated measures showing no interaction between age and user interface (F(1,72) =1.34, $p = .27$) for the transformed data.

According to our theoretical framework, the discontinuity around the age of 40 suggests that experience-related factors play a role here. As no person had ever used a videophone before, prior experience with such a device could not be an explanatory factor. Possibly, recent computer or mouse experience (before the experiment) could explain this generation effect. Table 3.1 shows the proportion of subjects with experience working with the computer and mouse per age group.

Computer experience diminished continuously with age and can not therefore explain the generation effect. Mouse experience, on the other hand, seems to suggest a discontinuity around the age of 50. As this trend does not correspond with the discontinuity between the age of 30 and 45 found for mode error performance, mouse experience is not expected to explain it either. A mixed repeated measures ANCOVA with mouse experience as the covariate confirmed this: the factor mouse experience was not significant (F(1,71) =1.00, $p = .32$), the main effect of age still was significant (F(3,71) =3.16, $p = .03$), and the Helmert Contrast analysis again showed a significant difference between the youngest group and the other three age groups (t= -2.80, $p = .01$). Hence, although the significant interaction between mouse experience and user interface (F(1,71) =7.88, $p =$ .01) indicated that mouse experience indeed helps subjects using the two-layered user interface, it did not explain the discontinuity.

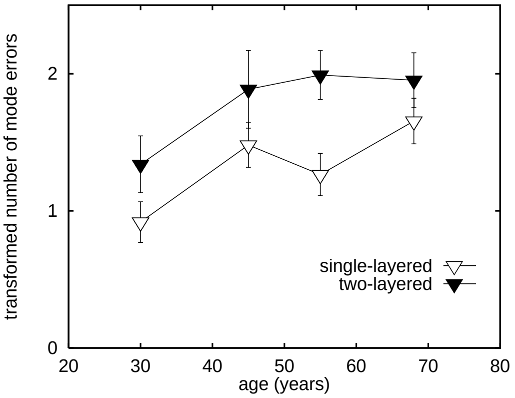


Figure 3.6: Experiment 1: Transformed mode error performance as a function of the average age per group and user interface. Each point reflects the average of twenty subjects. The vertical bars in this and the following figures denote twice the standard error of the mean.

Learning effects

Figure 3.7 shows transformed mode errors as a function of trial number and order of using single and two-layered interfaces. It can be seen that the four age groups roughly show the same pattern. Everyone was able to learn by repetition, including the older age groups. It is also visible that the youngest group shows most reduction of transformed mode errors in the second trial. This was confirmed by the Helmert contrast analysis (t= -2.26, $p = .03$). Considering the asymmetric change in results from trials 2 and 3, it can be observed that the order in which the

Table 3.1: Proportions computer and mouse experience per age group in experiment 1.

age category	computer	mouse
25-35	.90	.90
40-50	.75	.80
50-60	.65	.40
65-75	.40	.40

user interfaces were presented influences the error rate as well. The results on the ANOVA mixed repeated measures, mentioned in the beginning, confirmed these observations.

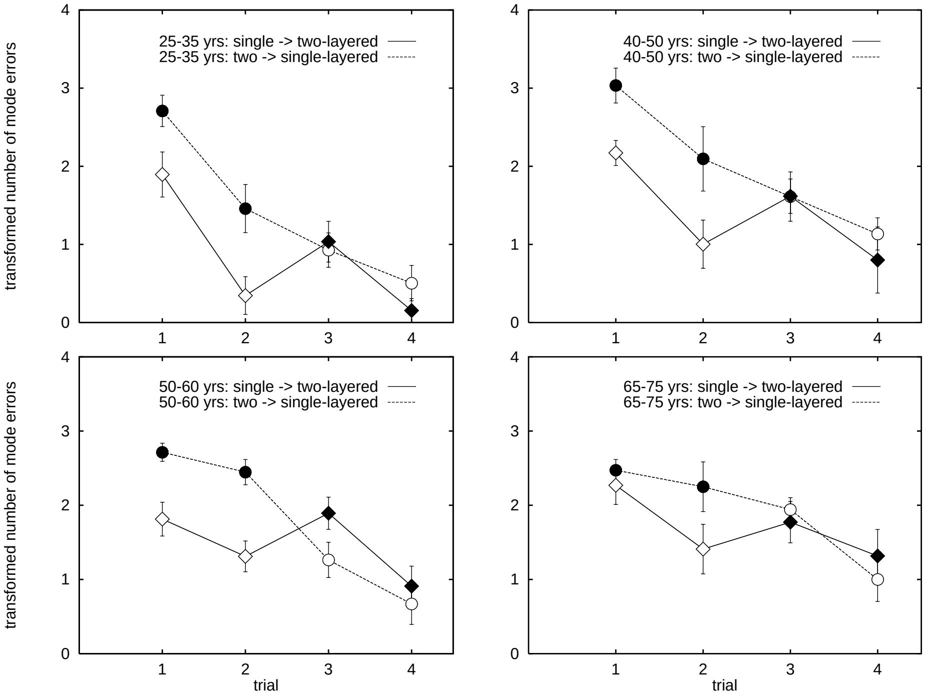


Figure 3.7: Experiment 1: Transformed error performance as a function of trial number and the order of using single or two-layered user interfaces, open and black symbols respectively. Each point reflects the average of twenty subjects.

Task duration analyses

The question is whether a similar result would be found when analysing the task duration. As is usual for time measures, the data were log-transformed [Winer, Brown and Michels, 1991]. Mauchly's Sphericity test indicated that the resulting variances were homogeneous (Mauchly's W =1.0). The 2 x 2 x 2 x 4 ANOVA mixed repeated measures with task duration per trial as the dependent measure showed that the main effects of age (F(3,72) =15.99, p = .00, power= 1.0), user interface (F(1,72) =73.26, p = .00, power= 1.0), and trial-per-user interface (F(1,72) =145.07, p = .00, power= 1.0) were significant. Moreover, the interaction of order by user interface (F(1,72) =91.02, p = .00, power= 1.0) and order by user interface by trial-per-user interface (F(1,72) =5.12, p = .03, power= .65) were found to be significant. Similar results were acquired by the untransformed data [Docampo Rama, 1997]. The following section explains the results on the transformed data in detail.

Age and generation

Figure 3.8 depicts the transformed task duration per user interface and age group. The age groups are presented by their mean age. The transformed task duration increased monotonically with age. The Helmert contrast analyses showed that not only the age group 25-35 differed from the older age groups (t= -5.78, p = .00), but also the group of 40-50 (t= -3.30, p = .00). The difference between the two oldest groups just failed to reach significance (t= -1.92, p = .06). For each of the user interfaces the monotonic increase can be described as an age effect.

According to our theoretical framework, neither computer nor mouse experience are expected to explain the age effect found, although the computer experience decreased monotonically with age, as well. A mixed repeated measures ANCOVA with computer experience as the covariate confirmed this. The analysis shows that the factor computer experience was significant (F(1,71) =10.36, p = .00) and an interaction of computer experience and user interface was significant (F(1,71) =4.92, p = .03) as well, but the age effect still remained (F(3,71) =10.31, p = .00). All subjects showed a shorter task duration by repetition and the learning effect also generalized towards the second user interface. This effect is in full agreement with the literature (e.g. Kelley [1996]).

3.3.4 *Discussion*

The results of experiment 1 indicate that there are age-group differences in using new software style appliances. Task duration showed a clear-cut age effect not related to generation. This is in line with other age-related findings in the literature (e.g. Salthouse, Hambrick, Lukas and Dell [1996]). On the other hand, the mode error performance across age groups was not related linearly or monotonically to age as found in the literature (e.g. Salthouse [1991]). It seems to be related to a generation effect as defined in our theoretical framework. The EM-generation (age of 40-75) had more difficulties using the videophone simulation than the S-generation (25-35 years). Moreover, no age-related differences in er-

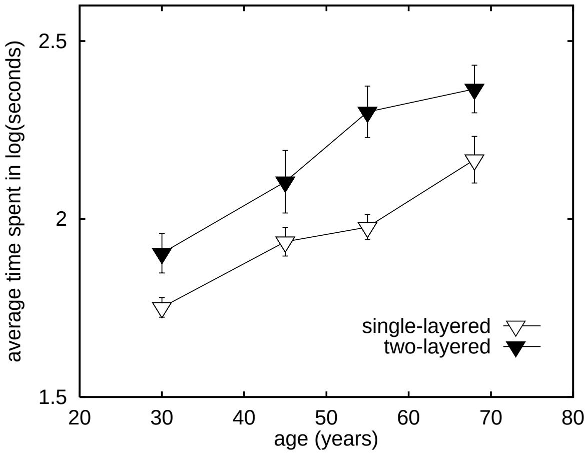


Figure 3.8: Experiment 1: Log-transformed task duration as a function of the average age per group for two types of user interfaces. Each point reflects the average of twenty subjects.

ror performance were detectable within the EM-generation.

In contrast to an age effect, a generation effect suggests no age-related cognitive factors to be involved, but experience-related aspects. The results suggest that recent experience with computer and mouse is somewhat helpful but cannot explain the generation effect. The technology generation theory, initiated by Sackmann and Weymann [1994], based on standard generation theories within the sociological sciences (e.g. Glenn [1974] and Cutler and Kaufman [1975]), further developed by Van de Goor and Becker [2001] and applied in this study for user interfaces, fits the generation effect found. Hence, generation-related experience with user interfaces in the formative period seems to influence the user's basic approach to new technology. Therefore, generation-related lack of knowledge about software style devices confronts the user above the age of 40 with performance difficulties using the new device.

Our results suggest that the type of effect found very much depends on the specific performance measure. Speed and error performance seem to measure different user aspects: the time measure seems to reflect the user's time management strategy in accordance with his/her cognitive capacities, whereas the number of mode errors measures the user's understanding of the system [Johnson, 1990] and therefore reflects the user's approach based on generational lack of interaction style experience. Note that the group of 40 to 50 therefore has a different speed-accuracy tradeoff than the older groups, as it has a speedy performance compared to the older age groups despite the similar generation-related difficulties using current devices.

In the literature, a generation effect in the use of complex devices has not been noted. This is probably because part of the studies use time measures to analyse age differences; see Salthouse [1991]. As has been found for task duration in this experiment, time measures only show an age effect. Other research mentioned in the literature only concern two age groups (e.g. Salthouse, Hambrick, Lukas and Dell [1996]).

Prior studies concerning more than two age groups that found differences in error performance in fact show a similar age group pattern as in this experiment, but this has not been referred to as a generation effect. Czaja and Sharit [1998] found that the average error rate while using a consumer electronics product with a software style was equal for the middle-aged (40-59 years) and the older subjects (60-75 years), but much higher than for the youngest group. An earlier study of Czaja [1996] shows a similar result. In that study she analysed the number of errors committed on a text editing task (again a software style device) during several types of training. A same number of errors was made by the middle-aged and by the older subjects, whereas fewer errors were committed by the youngest group. But, similar to our findings, the average time spent on the editing task in that study was purely age dependent.

The technology generation effect not only holds for the two-layered structure, but also for the single-layered one. This suggests that only changing the structure of software-based user interfaces into a single-layered structure visualizing all available functionality is not sufficient to compensate the generation-related lack of knowledge concerning current user interfaces. Detweiler, Hess and Ellis [1996] found that a grid display did not compensate for the age group differences either. Likewise, in the study of Gomez, Egan, Wheeler, Sharma and Gruchacz [1983] it appeared that a screen-based editor did not eliminate the age differences either.

Nevertheless, the single-layered interface was easier to use than the two-layered one. This is in line with the results found by Landauer and Nachbar [1985] concerning the advantage of broad, shallower menu trees compared to the narrow, deeper ones, and Detweiler, Hess and Ellis [1996] who compared the grid display with the window display. The effect of order even suggests that the single-layered structure has an overall facilitating role when using a new system. It helps the user to efficiently form a conceptual model of the system. Finally, the present experiment shows that all subjects are able to learn by practice, including older subjects.

3.4 Experiment 2: three-layered interface

In the second experiment we investigated whether the results from the first experiment hold for other, more complex systems. For this, a simulation of a mobile phone was run, featuring a three-layered user interface with two complex tasks.

In line with the results of experiment 1, a technology generation effect is expected for mode error performance, whereas an age effect is expected for task duration. Beside that, all age groups are expected to learn by repetition.

3.4.1 *Method*

Subjects

Forty subjects took part in the experiment that was carried out in 1998/1999. The age range was 25 to 75. The subjects were assigned to four age groups. The age of the youngest group ranged from 25 to 35 (representing the S-generation), with a mean age of 31. The age of the first middle group ranged from 40 to 50, with a mean age of 46, the second middle group ranged from 50 to 60, with a mean age of 55. The age of the oldest group ranged from 65 to 75, with a mean age of 69. The latter three age groups are part of the EM-generation.

Educational level was confined to high school diploma or bachelor degree (equally divided within each age group), resulting in a homogeneous experimental population in this respect. An equal number of female and male subjects was selected per age group to control possible gender differences. Finally, the subjects did not have any self-reported problems with manual movements, hearing or vision.

3.4.2 *Design*

The independent variables were the trial (first and second) and the age group (25-35, 40-50, 50-60, or 65-75). The experimental design was a 2 x 4 design. Age served as the between-subjects variable; trial was the within-subject variable. The dependent variables were the number of mode errors and task duration per trial.

Materials

One mobile phone interface was simulated using Microsoft Visual Basic 4.0 in the Windows 3.11 environment. Figure 3.9 depicts this user interface. Two complex tasks, a memory function and a function to change the colour of the cover, were implemented on the device representing part of the functionality of the simulation. The user interface was provided with buttons and a display. The buttons represented 10 digits, a star, a #, two function buttons, a confirmation button, two arrow buttons, and three memory buttons. Mode errors could be made by clicking on the confirmation, memory and arrow buttons in the wrong mode. All buttons and words were clearly and equally visible. The buttons could be activated by clicking with a mouse-pointing device. The experiment itself was run on a laptop Pentium 130. The subjects were seated in a quiet room that was well lit without specular light reflection on the laptop screen.

Procedure

Experiment 2 consisted of two trials. Again, no break was given between the trials. Within each trial, two tasks were executed in a fixed order. The participants were first asked to change the colour of the cover from red to blue. After that they were asked to program their own home telephone number under button C. Figure 3.9 depicts the shortest way to carry out these tasks. Both tasks consisted of pressing first a function button and subsequently a confirmation button. Direction buttons had to be used to change the colour of the cover. A memory button had to be pressed after entering a telephone number to save the telephone number. All subjects had to achieve the goal of each task before going on with the following task and trial. Nineteen steps were needed to carry out the tasks. Before starting the first trial, all subjects were given a general introduction and practiced with the mouse-pointing device in two exercises. After the final trial of the experiment, each subject was interviewed to obtain background information about the subjects, such as their computer and mobile phone experience.

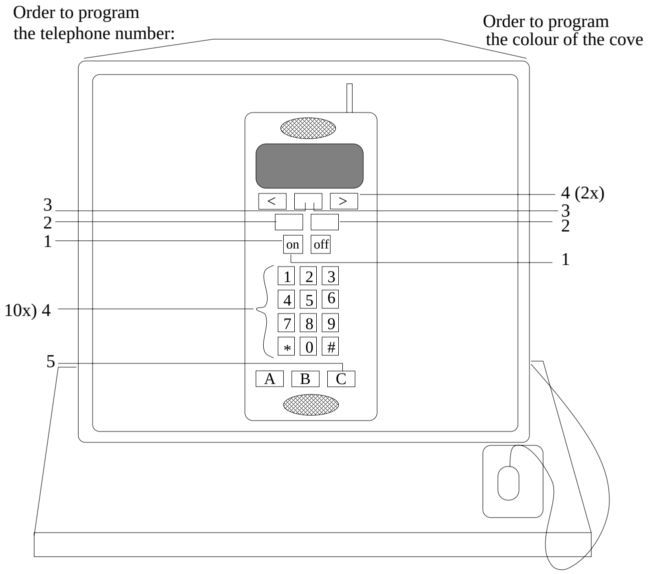


Figure 3.9: Experiment 2: Simulation of a mobile phone. The numbers denote the steps to be taken in sequence to carry out the task with minimal number of steps.

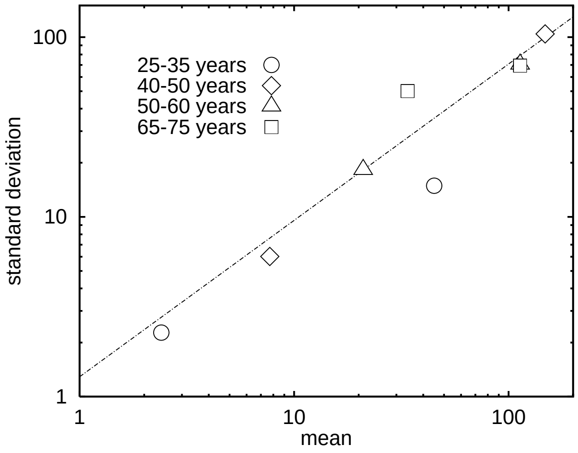


Figure 3.10: Experiment 2: Standard deviations as a function of the mean number of mode errors. Each point reflects the average of ten subjects.

3.4.3 *Results*

The optimal performance during the two trials was 2x19=38 steps. On average 172 steps were needed by the youngest group at the end of the two trials, whereas 418 steps[3] were made by the group of 40-50, 439 steps by the age group of 50-60 and 474 by the oldest group. The youngest group spent 6 minutes and 8 seconds on average to finish the experiment. The group of 40-50 years showed a total task duration of 19 minutes and 1 second. The age group of 50-60 finished in 20 minutes and 27 seconds on average, whereas the oldest group finished in 43 minutes and 15 seconds.

Mode error transformation

Each subject took part in 2 trials. Combined with the 4 age groups, this resulted in 8 conditions with 10 scores per condition. Again the mean and the standard deviation calculated for these conditions show that the standard deviation increases with the average number of mode errors (figure 3.10). The overall relation could be described by the following function:

$$\log s(x) = 0.11 + 0.87 \log \bar{x} (r = .88)$$

Hence, the standard deviation $s(x)$ is a power function of the mean $\bar{x}$ with the power of .87. Accordingly, for further analyses all raw scores were transformed

---

[3] The total number of steps needed corresponds with the sum of the minimal number of actions, the total number of mode errors and the total number of redundant steps needed by the subjects.

Table 3.2: Proportion computer and mouse experience per age group in experiment 2.

age category	computer	mouse
25-35	1.00	.80
40-50	.80	.80
50-60	.70	.60
65-75	.60	.60

by raising them to the power of .13. Only the transformed data were used for the analyses.

Analyses of mode errors

An ANOVA repeated measures was carried out with the transformed number of mode errors per trial as the dependent variable. In this experiment the main effect of age (F(3,36) =4.68, $p = .01$, power= .87) and trial (F(1,36) =99.73, $p = .00$, power= 1.0) were found to be significant. No interaction between age and trial was found (F(3,36) =1.73, $p = .18$, power= .11). The following sections explain these results in detail.

Age and generation

Figure 3.11 depicts the error performance per age group. The four age groups are represented by their mean age. The figure shows that the youngest age group made fewer mode errors than the other three age groups. The three oldest groups did not show an age-related increase of transformed mode errors. As mentioned before, such discontinuity can be interpreted as a generation effect, which again was found between the age of 30 and 45. The Helmert Contrast analyses confirmed this effect. Only the age group of 25-35 differed significantly from the three older age groups (t= -3.71, $p = .00$). The age groups from the age of 45 on showed a similar error performance (t= -.19, $p = .85$ for the difference of the group of 40-50 years versus the older groups, and t= .49, $p = .63$ between the oldest groups). This discontinuity suggests that experience-related factors play a role here.

Surprisingly, no subject had ever used a mobile phone before. Therefore, prior experience with such a device could not be an explanatory factor. Can recent computer or mouse experience (before the experiment) explain this generation effect? Table 3.2 shows the proportion of subjects with mouse and computer experience per age group. In this experiment computer experience declined with age too and can therefore not explain the generation effect. Again, mouse experience showed a discontinuity around the age of 50. This trend does not cor-

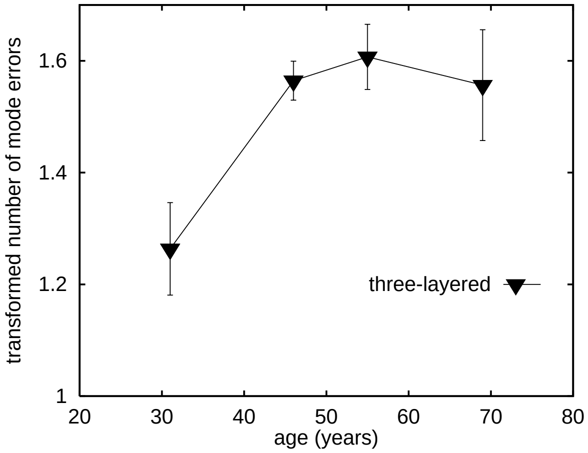


Figure 3.11: Experiment 2: Transformed mode error performance using a three-layered user interface as a function of the average age in the four groups. Each point reflects the average of ten subjects.

respond with the discontinuity found for mode error performance. A repeated measures ANCOVA with mouse experience as the covariate confirmed this. The factor mouse experience was not significant (F(1,35) = .21, p = .65), the main effect of age remained (F(3,35) =4.35, p = .01), and the Helmert Contrast analyses again showed a significant difference between the youngest group and the older age groups (t= -3.58, p = .00). So far, the effect only fits the technology generation effect described in section 3.2.3 that differentiates generations of users by their interaction style experience in the formative period.

Learning effect

Figure 3.12 depicts the transformed error performance per age group and trial. The age groups are represented by their mean age. No significant interaction between age and trial was found, which indicates that the transformed number of mode errors is reduced at the same amount for all age groups. It shows that everyone is able to learn by repetition, irrespective of age. Note that the generation effect remains during the second trial.

Task duration analyses

The task duration data were log-transformed. Mauchly's Sphericity test indicated that the variances of the transformed data were homogeneous (Mauchly's W =1.0). The 2 x 4 ANOVA repeated measures with task duration per trial as the

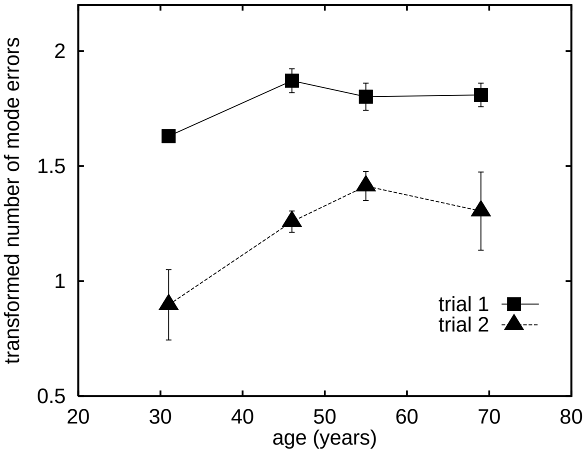


Figure 3.12: Experiment 2: Transformed mode error performance using a three-layered user interface depicted per average age within each group and trial number. Each point reflects the average of ten subjects.

dependent measure showed that the main effects of age (F(3,36) =15.42, $p = .00$, power= .87) and trial (F(1,36) =197.33, $p = .00$, power= 1.0) were significant. The following section explains these results further.

Age and generation

Figure 3.13 depicts the transformed task duration per age group. The age groups are presented by their mean age on the age scale. In this experiment the task duration increases monotonically with age. The Helmert contrast analyses not only showed that the age group 25-35 differs from the other age groups (t= -6.07, $p = .00$), but from the group of 40-50 (t= -2.64, $p = .01$) too. The difference between the two oldest groups failed to reach significance (t= -1.57, $p = .13$). The effect found can be described as an age effect, which may be related to cognitive changes with age. Hence, neither computer nor mouse experience are expected to explain the age effect found, although the computer experience decreased monotonically with age. A repeated measures ANCOVA with computer experience as the covariate confirmed this. The analysis showed that the factor computer experience indeed does not play a role (F(1,35) =2.84, $p = .10$) and the age effect still remains (F(3,35) =11.70, $p = .00$).

Not only fewer mode errors, but also a shorter task duration was seen during the second trial, irrespective of the age group. The age effect persisted in the second trial as well.

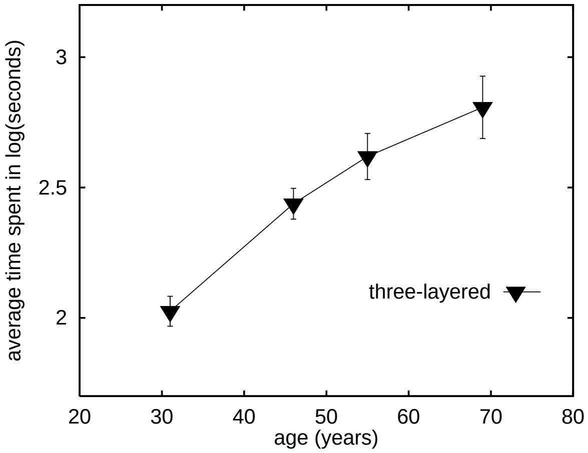


Figure 3.13: Experiment 2: Task duration as a function of the average age within each group. Each point reflects the average of ten subjects.

3.4.4 *Discussion*

The results of study 2 indicate age group differences using a three-layered user interface. The age group differences in mode error performance show a similar discontinuity between the age of 30 and 45 as in the first experiment, suggesting a similar generation effect. In other words, only the youngest age group, part of the S-generation differs significantly from the other age groups that are part of the EM-generation. No age-related increase of mode errors has been found within the EM-generation. Hence, age-related cognitive changes as found in the literature (e.g. Salthouse [1991]) do not resemble the error performance across age groups. The mode error analyses suggest that recent experience with modern devices such as the mobile phone, computer and mouse does not explain it either. Generational differences in experience with the software style as identified in Docampo Rama and Van der Kaaden [2001] appear to correspond with the discontinuous trend in error performance found. The task duration analyses showed a similar age effect as found in experiment 1 and the age effects mentioned in the literature. These results reinforce the idea that the effects depend on the type of measure used to analyse user performance.

Thus, in line with research of Czaja [1996] and Czaja and Sharit [1998] the two experiments mentioned in this chapter, error performance and task duration, show a different trend. According to the explanation given in the discussion of experiment 1, this indicates that the two measures reflect different user aspects that persist over tasks, user interface complexity and educational level. In the second experiment the speed-accuracy tradeoff of the group of 40-50 remains different from older subjects that are part of the same technology generation.

According to the number of errors committed, the number of steps needed and the time spent, it can be said that the user interface was found to be rather complex, which means that a considerable cognitive effort had to be made. This supports the result found by Landauer and Nachbar [1985] concerning the disadvantage of narrow, deeper menu trees. Similar to earlier findings, exercise reduces the error rate and the time needed to carry out the task, irrespective of age. However, the generation effect found for mode errors does not disappear in the second trial nor does the age effect found for task duration.

3.5 General conclusion

Three main effects have been found in this study: (1) a generation effect was found for error performance, whereas an age effect was found for task duration, (2) the deeper the interaction structure, the more difficult it is to use a complex system, and (3) all users are able to learn by repetition, and this learning effect is also visible when working subsequently with a different user interface of the same product.

Thus, it depends on the type of measures used whether age group differences can be described as age or generation effects. Task duration mainly measures the user's time management strategy in accordance with his/her cognitive capacities, whereas mode error performance reflects the user's misunderstanding of the system. The error performance analyses indicated that current user interfaces are better understood by users that are members of the software generation, than by those who are members of the electro-mechanical generation. Experiencing modern user interfaces in the formative period seems to be an advantage when using a new device, irrespective of complexity, tasks and educational level. Finally, a present-day user interface with a single-layered structure as seen in this study is easier to use and facilitates learning.

3.5.1 *Design implications*

Compared to a user interface with a software style provided with a multi-layered structure, a user interface with a single-layered structure is easier to learn and has an overall facilitating role during the user's learning process. Hence, an overview of the functionality available, direct feedback about the status of the device, and a direct link between the control and its function are useful features during the learning process of all users, but only within the user's generation limitations. While task repetition helps the user to reduce the error rate, features like recognizable feed forward and on-line learning instructions [Masthoff, 1997] may help inexperienced users.

3.5.2 *Implications for further research*

It would be very appropriate to analyse which factors underly the age and generation effects found. For this, cognitive abilities should be measured explicitly next to technology experience in the formative period and later in life. Moreover, as the speed-accuracy tradeoff differs per age group, it is expected that user strategies resemble a mix of age and generation effects.

4

The role of visuo-cognitive skills in handling layered interfaces

Particularly older users have difficulties using present-day user interfaces. In an previous chapter a continuous increase with age was observed in task duration indicating an age effect, but a discontinuous increase in error performance pointing to a generation effect . This chapter analyses how these effects are related to cognitive abilities, in particular visuo-spatial working memory and reasoning ability. Performance of 88 subjects between 25 and 80 years was measured using a simulated mobile phone. The cognitive abilities were measured by two selected neuropsychological tests. The generation effect in error performance was related to visuo-spatial working memory in relation to advanced hand-eye coordination that seems to be better trained by the software generation (born in 1960 and later) than by the electro-mechanical generation (born before 1960). Within this older generation a speedy trial-and-error strategy turns out to establish a larger learning effect than an accurate and cautious, reflective way of proceeding.

4.1 Introduction

Older users face more problems with present-day consumer electronics products than younger users [Stewart, 1992; Kelley, 1996; Freudenthal, A., 1998]. Current user interfaces can be characterized by its multiple specialized functionality, software-based input and output devices, a central display on which information pops-up, its multi-layered structure, programming and navigational procedures. This interface architecture does not support a one-to-one relation between the function and the button, the multiple functions of buttons depending on their mode, and lacks adequate feedback or feed forward. In a historical study [Docampo Rama and Van der Kaaden, 2001] this type of user interface has been labeled the *software style*.

Authors of user-system interaction studies related age-dependent problems with present-day user interfaces to the user's motor and cognitive abilities. Salthouse, Hambrick, Lukas and Dell [1996] indicated that the user's computer task efficiency was inversely correlated with his/her reaction time on motor tasks. Next to reaction time, Freudenthal, T.D. [1998] found that the user's latency when working with an interactive system was negatively related with spatial memory and reasoning abilities. Also Czaja and Sharit [1998] found that visuo-motor skills and memory correlated with the user's error performance on data entry tasks.

Several cognitive abilities tend to decline monotonically with age in adults [Salthouse, 1991; Birren and Schaie, 1990; Craik and Bosman, 1992]. In our previous study we also found this trend for task duration. However, error performance showed a discontinuous increase with age (see the left-hand panel of figure 4.1). Taking a closer look in studies of Czaja [1996] and Czaja and Sharit [1998] and plotting them next to our previous results we found a similar discontinuous trend in error performance (see the right-hand panel of figure 4.1). In the study of Czaja [1996], next to a discontinuous trend in error performance, also a continuously increasing trend with age could be observed for task duration.

The discontinuous trend found in our prior study (described in chapter 3) could be related to the user's technology experience earlier in life. People often rely on their most established knowledge, which is mostly learned during their adolescence [Rubin, Rahhal and Poon, 1998]. Child development studies indicate that in this same period of life cognitive abilities and experiences influence each other strongly for further maturation [Sroufe and Cooper, 1988]. But also in sociological studies [Glenn, 1974; Glenn, 1977; Cutler and Kaufman, 1975] the period between the age of 10 and 25 has been indicated as a critical period in which people acquire norms, values and skills that influence future behaviour. This period has been labeled as the *formative period*. Based on this, several generations can be distinguished of subsequent birth-cohorts that encountered a sim-

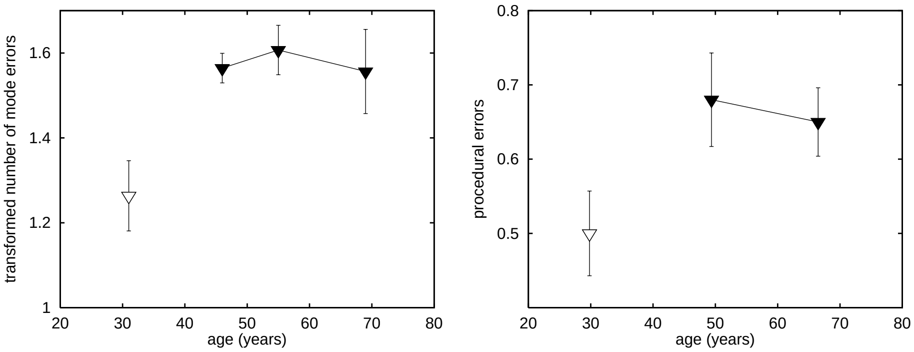


Figure 4.1: Left-hand panel: Transformed number of mode errors per age group using a mobile phone simulation as found in our previous study . Right-hand panel: Procedural error performance using a data entry task in a study of Czaja and Sharit [1998]. The vertical bars in these and following figures denote twice the standard error of the mean. In this and the following figures, the open symbols represent the scores of the software generation, whereas the closed symbols represent those of the electro-mechanical generation.

ilar environment in the formative period and display similar behaviour nowadays. The type of environment in the formative period has to be stable for a long enough period to be able to identify a specific generation. Current behaviour as effected by being a member of a certain generation can be distinguished from a traditional age effect by its discontinuous trend instead of a continuous one [Glenn, 1977]. That is why the results depicted in figure 4.1 can be interpreted as generation effects.

Sackmann and Weymann [1994], and Van de Goor and Becker [2001] analysed generation effects considering specifically the user's technology experience. Results of Sackmann and Weymann [1994] confirmed that people who experienced the availability of the same types of domestic products during the formative years, display currently similar purchase behaviour. Based on these results they defined several technology generations. Also Van de Goor and Becker [2001] found that different technology generations differ in their possession and use of current appliances.

In line with these studies, it was investigated in our previous study (chapter 3) whether present-day Information, Communication, and Entertainment devices (abbreviated as ICE devices) are more easily learned by the software generation (the only ones who acquired software style experience in their formative years) than by older generations. If so, the discontinuity in the performance as a function of age would be expected around the birth cohort of 1960, as earlier birth cohorts did not encounter the software style in their formative years [Docampo Rama and

Van der Kaaden, 2001]. Both age and technology generation effects were found, depending on the type of performance measure that we focused on. An age effect was visible for task duration performance, whereas a generation effect was seen for error performance. These results were found for all software style devices that were studied, irrespective of its number of layers and for two educational levels.

In this study we investigate age and generation trends in performance across and within technology generations. In particular, these trends will be related to specific cognitive and motor abilities. The next sections explain which factors we will focus on.

4.1.1 *Cognitive abilities*

Visuo-spatial working memory

Working memory can be described as a temporary storage in which incoming information, acquired from the environment and the person's long term memory, is kept available for the completion of the task [Baddeley and Hitch, 1974]. It contains slave systems to hold and rehearse information and a central executive that regulates further processing [Baddeley, 1986]. One of the slave systems is the visuo-spatial sketchpad, a temporal store for visual and spatial information [Logie, 1995].

Age-related working memory problems arise in tasks where subjects hold information while carrying out rather complicated processing activities [Craik, Morris and Gick, 1990; Craik and Bosman, 1992; Salthouse and Skovronek, 1992; Howard and Howard, 1997]. Age-related difficulties during simultaneously holding, processing, and manipulating material show up when combined storage and processing requirements exceed the capacities of working memory [Salthouse, 1991]. The age decline is most pronounced for working memory tasks that address visuo-spatial information [Salthouse, 1991; Hale, Bronik and Fry, 1997].

Human-computer studies suggest that all users need their visuo-spatial working memory to interpret, hold, and process visual and spatial information on a user interface to activate a function of the equipment [Gomez, Egan, Wheeler, Sharma and Gruchacz, 1983; Detweiler, Hess and Ellis, 1996]. Age group differences in using present-day ICE devices have therefore been related to the age-related decrease in visuo-spatial working memory [Czaja, 1996; Czaja and Sharit, 1998].

Many visuo-spatial working memory tests can be used to check the user's ability in this respect. For our study we have been looking for a test that emphasizes the memory of visuo-spatial information while performing. Therefore, the Corsi block-tapping test will be used. It is a visuo-spatial working memory test that resembles the way input devices, such as buttons, have to be controlled by the user when working with a user interface. It emphasizes the user's memory of the spatial ordening of the tapped objects on a two-dimensional display.

Reasoning ability for visuo-spatial information

Next to visuo-spatial working memory, reasoning ability for visual and spatial information has been mentioned to play a role when using complex systems [Reason, 1990; Freudenthal, T.D., 1998]. Normative studies of the Raven's Matrices test [Raven, 1960] show that abstract reasoning with visual material declines almost linearly with age starting from about the age of thirty onwards [Salthouse, 1991].

As the Raven's Matrices test measures the user's abstract analysis for visual material, without having to memorize it, in this study this test will be measured next to visuo-spatial working memory.

4.1.2 *Motor abilities*

Movement ability

Salthouse, Hambrick, Lukas and Dell [1996] found that the user's time performance on computer tasks was explained for 70 percent with movement duration measures of reaction time tasks. In this paper, the user's shortest movement duration between two objects on a two-dimensional display [Fitts, 1964] will be measured as a possible contributing factor of the user's performance while dealing with a device.

Speed-accuracy tradeoff

Older and younger subjects differ in what is called their speed-accuracy tradeoff [Welford, 1958; Welford, 1968]. Studies described by Welford indicate that during complex tasks older users tend to lengthen their movement duration per action to make careful decisions. Hence, older users prefer to be cautious and accurate (reflective strategy) instead of fast and erroneous (speedy trial-and-error strategy). Such a choice of older users is in fact a coping strategy to maximize their overall achievement while performing complex tasks.

In the current experiment subjects will neither be asked to be as fast as possible, nor as accurate as possible. The only instruction they get is to reach the goal. In such setting subjects have the opportunity to choose their own balance between speed and accuracy. Therefore, in this study the user's average duration per action during the experiment and the total number of errors will be plotted to capture the user's strategy in this respect.

4.1.3 *Technology experience*

Historical analyses [Docampo Rama and Van der Kaaden, 2001] have shown that two main types of user interfaces or interaction styles have been implemented on ICE-devices in the past century: the electro-mechanical style (approximately from 1930 to 1980) and the software style (from 1980 on). This means that the generation of people that were born before 1960, grew up with the first style only, and can therefore be named the electro-mechanical generation. The generation born in 1960 or later is called the software generation. In Docampo Rama and Van der Kaaden [2001] the boundary between these generations has been established by asking 1015 subjects whether they used a wireless remote control in their formative period. Survey analyses showed that this marker was indeed highly correlated with the boundary between the electro-mechanical and software generation as found by historical analyses. Therefore, this marker is also used in this paper to verify the technology generation boundary. The boundary between the electro-mechanical generation and the software generation is again expected to be associated with a discontinuous trend in error performance.

Next to generation-related experience, other technology experience questions will be asked to catch the importance of current software style experience when using present-day ICE devices.

4.1.4 *Present research*

The objective of the present study is to determine how the continuous trend with age in task duration and the discontinuous trend with age in error performance are related to visuo-spatial working memory and reasoning capability with visual material. Also the user's movement ability, speed-accuracy tradeoff and technology experience will be considered. To this end the experiment has been carried out on subjects between 25 and 80 years using a mobile phone simulation.

4.2 Method

4.2.1 *Subjects*

The experiment was executed in 1999. Eighty-eight subjects took part in the experiment. The age range was 25 to 80 with 8 subjects per age cohort of 5 years, leading to 11 age groups.

In line with the study of Docampo Rama and Van der Kaaden [2001], the subjects were asked whether they had experience with a wireless remote control before the age of 25. Also in the current study 96 percent of the subjects born in 1960 and later, which were 39 and younger (first three age groups), were acquainted with the wireless remote control before the age of 25. Those born before 1960 had not experienced the wireless remote control in their formative years. Therefore, the boundary between the electro-mechanical and software generation will be set again around 1960 as year of birth.

The educational level was confined to high school diploma or bachelor de-

Table 4.1: Number of subjects out of eight per age group that had ever controlled a mobile phone.

age group	N
25-29	5
30-34	6
35-39	4
40-44	3
45-49	5
50-54	3
55-59	2
60-64	1
65-69	2
70-74	3
75-79	0

gree (equally divided within each age group), resulting in a rather homogeneous experimental population in this respect. Four female and four male subjects were selected per age group to control for possible gender differences. The subjects did not have any self-reported problems with manual movements, hearing, or vision.

As a simulation of a mobile phone was used, it was asked whether the subjects had used a mobile phone before. Part of the subjects within each age group had at least called once with a mobile phone before, except for the oldest age group. Table 4.1 shows per age group the number of subjects with mobile phone experience. Furthermore, it was enquired how many software-style devices they had used until now. The devices mentioned in this respect were the television with remote control, teletext, video recorder, compact-disc player, wireless telephone, mobile phone, fax, computer, cd-rom, and Internet. Members of the software generation have used on average eight different software-style devices until now. Within the electro-mechanical generation the number of software-style devices decreased with age to around three devices. Figure 4.2 depicts this trend.

4.2.2 *Materials*

A simulation of a mobile phone was run, featuring a three-layered user interface, to measure the user's task duration, error performance, and speed-accuracy tradeoff. Furthermore, specific neuropsychological tests were used to assess the user's visuo-spatial working memory and reasoning ability for visual material. These materials were operationalized as follows.

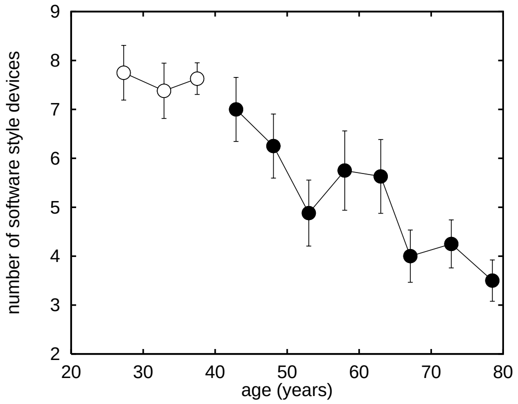


Figure 4.2: Average number of software style devices per age group concerning the number of different software style appliances that were used until now. Each point reflects the average of ten subjects.

The simulation

A mobile phone interface was simulated using Microsoft Visual Basic 4.0 in the Windows 3.11 environment. Figure 4.3 depicts this user interface. Total task duration, number of steps needed, and number of mode errors per trial were logged. Here a mode error is a redundant action of the user in which a button press does not lead to a system reaction as it has no function in that particular mode.

Two tasks, programming a telephone number and changing the colour of the cover, were implemented on the device representing part of the functionality of the simulation. The user interface was provided with a display and 22 buttons: 10 digits, a star, a #, an on-button, an off-button, two function buttons, a confirmation button, two arrow buttons, and three memory buttons. Mode errors can be made by clicking on the confirmation, memory, and arrow buttons in the wrong mode. All buttons and words were large and equally visible. The buttons could be activated by clicking with a mouse-pointing device. Performance measures per trial were the *total number of mode errors*, *total task duration*, *average duration per step*, and *total number of steps*. The experiment itself was run on a laptop Pentium 130. The subjects were seated in a quiet room that was well lit without specular light reflection on the laptop screen.

Mouse-pointing device exercise

Two mouse-pointing exercises were given. The first exercise concerned tapping a ball depicted on the screen that jumped away after a click. After ten subsequent

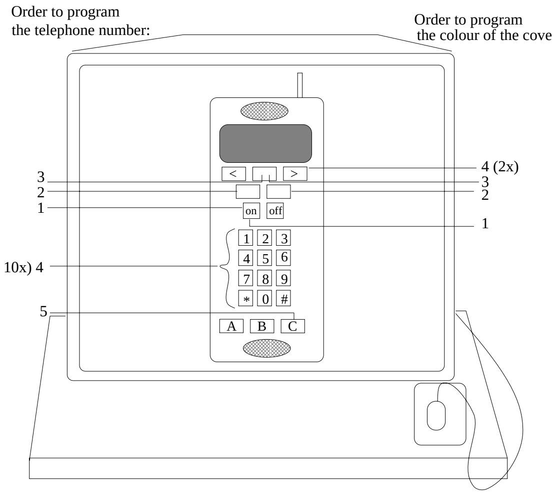


Figure 4.3: Simulation of a mobile phone as used in the experiment. The numbers denote the steps to be taken in sequence to carry out the task with a minimal number of steps.

clicks on the ball, the second exercise started. The task concerned clicking 10 times as fast as possible on two separate squares. The shortest time span while moving between the two squares was counted as the user's *shortest movement duration* that is related to the user's overall processing speed.

Progressive matrices test of Raven

The progressive Matrices test of Raven [1960] is a neuropsychological test that has originally been used as a measure for intelligence (IQ). It measures the subject's problem solving or *reasoning ability* when confronted with the presented visual material. The theoretical assumption behind it is that during such a test subjects use working memory for visuo-spatial information to find out the solution by abstract reasoning without the need to memorize the material. The test is divided in five sections with puzzles that gradually increase in difficulty. A total of sixty reasoning problems have to be solved, the first two of which are exercises. Every problem consists of completing a pattern. Six example patterns were presented to the subjects below the incomplete pattern. Only one of these example patterns is the correct answer (multiple-choice). The highest score is sixty, when all patterns are filled in correctly. The subjects carried out this test without time restriction.

Corsi's block-tapping test

Corsi's Block-Tapping Test [Corsi, 1972], as modified by Milner [1971], measures the user's *visuo-spatial working memory ability*, i.e. the ability to memorize or hold visuo-spatial information in working memory while manipulating objects in a certain order. In this test nine blocks are placed on a board in a certain a-symmetric pattern. The experimenter taps some of these blocks and the subject is asked to tap the same blocks in the same order. The number of tapped blocks increases in the course of the test. The original test is most difficult for subjects with visuo-spatial working memory disabilities. In the present study a computerized, alternative version of the test was used to relate it more closely to visuo-spatial working memory problems when using complex software-based devices. In this automated version of the Corsi's block-tapping test, the blocks were placed in rows and columns, forming a symmetric pattern. Furthermore, a red dot appeared on the blocks that had to be remembered by the subject, and the experimenter did not have to tap. Finally, the subjects did not tap the blocks, but were asked to click on the blocks in the same order by means of a mouse-pointing device. The tapping sequences remained the same as in the original test.

The test started with an exercise in which only two blocks were marked by the red dot. This exercise stopped when the subject was able to tap these two blocks in the correct order. After this, the test starts. Subjects have to remember and repeat the order of the tapped blocks. The test started with two blocks to be remembered, then three, then four, etc. These were the different rounds. In each round, the subject got two runs. The De Renzi's procedure [De Renzi, Faglioni and Previdi, 1977] was used for scoring: if the right blocks are tapped in the correct order both of the times, the subject gets two points in that round. If only one of the times the right blocks are tapped, only one point is given. If the participant makes a mistake on both runs of the same round, the test stops. The maximum score is eighteen.

4.2.3 *Procedure*

The mobile phone experiment consisted of two trials. No break was given between the trials. Within each trial, two tasks were executed in a fixed order. The subjects were first asked to change the colour of the cover from red to blue. Subsequently, they were asked to program their own home telephone number under button C. Figure 4.3 depicts the shortest way to carry out these tasks. Both tasks consist of pressing first a function button and after that a confirmation button. Arrow buttons had to be used to change the colour of the cover. A memory button had to be pressed after entering a telephone number to save it. All subjects had to achieve the goal of each task before going on with the following task and trial. Nineteen steps were needed at minimum to achieve the two mobile phone tasks. Before starting the experimental trials, all subjects were given a general instruction and exercised with the mouse-pointing device twice.

After the mobile phone experiment, the computerized version of the Corsi block-tapping test was taken. Subsequently, the subjects were seated in a different room and were interviewed about their technology experience. Finally, the progressive matrices test of Raven was given. A break and instructions were given before each test. The total experiment took on average one hour and thirty minutes.

4.3 Results

4.3.1 *Cognitive factors*

The main question of this paper is whether cognitive factors like reasoning ability and visuo-spatial working memory can be related to mode errors and task duration performance across and within technology generations while handling a software style device. For this, we first need to consider how the user's cognitive abilities vary as a function of age.

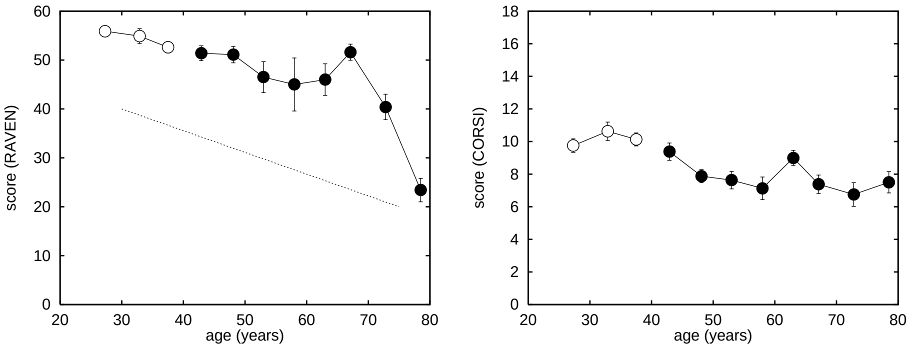


Figure 4.4: Cognitive abilities per age group. Left-hand panel: Average reasoning ability scores for visual material (Raven's test) per age group. The dotted line represents the Raven's Matrices test scores with age taking the average of four normative studies [Vernon, 1947, Slater 1947, Heron and Chown, 1967, and Burke, 1985]. Right-hand panel: Mean of visuo-spatial working memory scores per age group. Each point in these and following figures represents the average of eight subjects. An open symbol represents scores of the software generation, and a black symbol those of the electro-mechanical generation.

Reasoning ability: The left-hand panel of figure 4.4 shows that the reasoning ability scores decline monotonically with age until the age of 70. From around the age of 70 a sharp decline is visible. Moreover, the spread is rather small, suggesting little variation between individuals. The dotted line within the figure depicts the average trend of four normative studies done by Vernon [1947], Slater [1947], Heron and Chown [1967], and Burke [1985]. The reasoning ability scores in the present study are substantially higher than the average of these four normative studies. This is probably due to the high educational level of the subjects who participated in this study. Results of Vernon [1947] seem to confirm this suggestion as she found a high correlation between scores in the upper ranges of the test and the occupational level. The figure shows also that subjects beyond the age of 75 produce low reasoning ability scores compared to the other age groups with a similar educational background. This is in line with Raven [1960], who mentioned that the Raven's Matrices Test is not appropriate for persons beyond the age of 75 as they would fail on this test. Therefore, normative studies have only focused on people younger than 75. Note however, that the present group beyond 75 years still shows on average a higher score than the oldest age groups of the normative studies.

Visuo-spatial working memory: The right-hand panel of figure 4.4 shows that Corsi's visuo-spatial working memory scores decline with age as well. Although, the effect is rather small. No normative studies have been done on the Millner's version of the Corsi block-tapping test to compare with [Bouma, Lindeboom and Van Houte, 1990]. However, according to Corkin [1983], a score between 5 and 9 is within the norms of older subjects.

4.3.2 *Error performance*

The left-hand panel of figure 4.5 shows the average number of errors per age group and session committed during the experiment plotted against their variance. As can be seen the mode error performance does not show a homogeneous variance. The overall relation can be described by the following function:

$$\log s(x) = 0.06 + 0.91 \log \bar{x} \quad (r = .95)$$

Hence, the standard deviation ($s(x)$) is a power function of the mean ($\bar{x}$) with the power of 0.91. For further analyses like repeated measures ANCOVA and regression analysis, the standard deviation should be independent of the mean number of mode errors. This can be accomplished by raising the raw data to the power of 0.09, i.e. 1 - 0.91; see Winer, Brown and Michels [1991] for a detailed explanation. This transformation shows a slightly different power than in chapter 3, due to the fact that it concerns the data of different subjects than those who participated in previous experiment. The right-hand panel of figure 4.5 shows the transformed error performance per birth cohort of 5 years for the two trials separately.

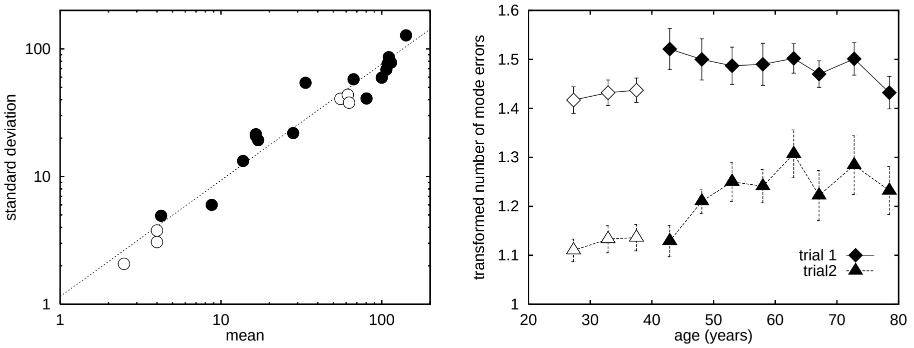


Figure 4.5: Error performance. Left-hand panel: Standard deviations as a function of the mean number of mode errors. Right-hand panel: Transformed number of mode errors per age group. The vertical bars in this and the following figures denote twice the standard error of the mean.

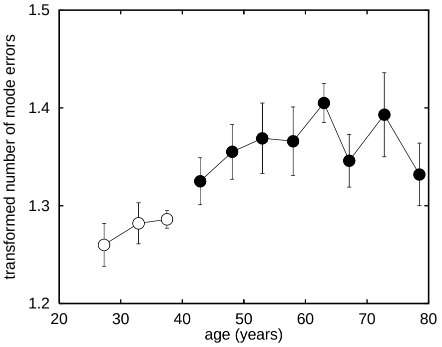


Figure 4.6: Transformed number of averaged mode errors across trials depicted per age group.

A repeated measures ANCOVA has been done with mode error as the dependent variable, trial as the within variable and age as the covariate. The analysis indicated that there is a significant effect of trial (F(1,86) =74.483, p= .000, power= 1.0), age (F(1,86) =15.060, p= .000, power = 1.0), and interaction between trial and age (F(1,86) =10.055, p= .002, power= .91).

The right-hand panel of figure 4.5 depicts the two trials separately per age group. In accordance with the analysis, the figure depicts that all age groups made fewer number of mode errors in the second trial than in the first trial, which is an indication that they learned.

Age and technology generation

The age effect as indicated by the repeated measures ANCOVA suggests a relation with age and/or technology generation. A stepwise regression analysis with the transformed number of averaged mode errors across trials as the dependent variable and age and technology generation as the independent variables showed the best fit for technology generation (F(1,86) =18.671, p= .000, power= 1.0). Figure 4.6 depicts the transformed number of averaged mode error performance across trials.

The regression analysis on the software generation indicated no significant correlation with age for the transformed number of averaged mode errors across trials (F(1,20) = 0.201, p= .659, power= .03). Neither did the regression analysis on the electro-mechanical generation indicate a significant correlation with age (F(1,64) = 0.857, p= .358, power= .08). Hence, the level of error performance of each technology generation seems to consist of a plateau.

The trend of the transformed number of averaged mode errors across trials (figure 4.6) was analysed further in relation to the user's cognitive abilities.

Average error performance

Scores of visuo-spatial working memory and reasoning ability were entered in a stepwise regression analysis with the transformed number of averaged mode errors as the dependent variable. The regression analysis shows that the correlation of error score with visuo-spatial working memory is significant (F(1,86) =14.435, p= .000, power= 1.0).

A similar analysis was done within each generation. A stepwise regression analysis on the software generation, with visuo-spatial working memory and reasoning ability as the independent variables, indicated a significant correlation with reasoning ability (F(1,20) =11.016, p= .003, power= .87). A stepwise regression analysis on the electro-mechanical generation indicated that none of the two cognitive variables correlated significantly with the error score.

4.3.3 *Task duration performance*

The task duration per age group and trial did not show a homogeneous variance either (left-hand panel of figure 4.7). As is usual for time measures, the data were log-transformed [Winer, Brown and Michels, 1991]. The right-hand panel of figure 4.7 shows the transformed task duration data per birth cohort of 5 years for both trials separately.

A repeated measures ANCOVA was carried out to analyse with the transformed task duration as the dependent variable, trial as the within variable and age as the covariate. A significant effect of trial was found (F(1,86) =82.129, p= .000, power= 1.0), but also of age (F(1,86) =104.420, p= .000, power= 1.0).

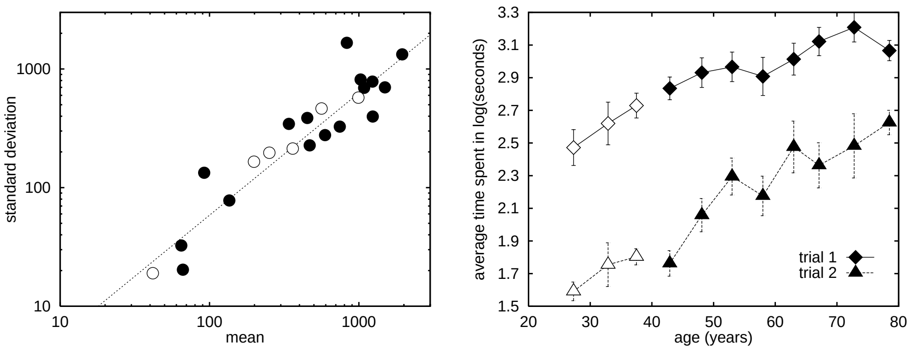


Figure 4.7: Standard deviation in function of the mean task duration.

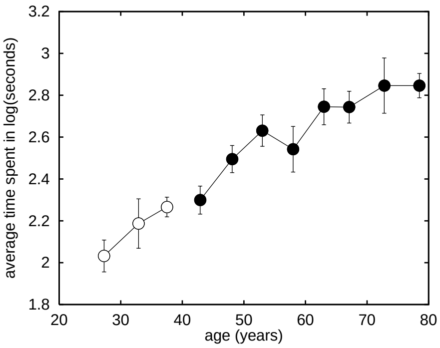


Figure 4.8: Left-hand panel: Average log-transformed task duration depicted per age group.

The interaction between age and trial indicated to be significant as well (F(1,86)= 2.054, p= .016, power= .68).

The right-hand panel of figure 4.7 depicts the log-transformed task duration per age group and trial. It confirms the results of the repeated measures ANCOVA: all age groups are able to learn.

Age and technology generation

Figure 4.8 depicts the log-transformed averaged task duration across trials. As confirmed by the ANCOVA repeated measures mentioned before, it shows an increase with age.

A stepwise regression analysis with the log-transformed averaged duration as the dependent variable and age and technology generations as the independent variables showed that the fit of age gave a higher correlation (F(1,86) =104.420, p= .000, power= 1.0), The log-transformed averaged task duration and reduction of log-transformed duration were analysed further in relation to the user's cognitive abilities.

Average task duration

A stepwise regression analysis with the log-transformed averaged task duration as the dependent variable and visuo-spatial working memory and reasoning ability as the independent variables showed that firstly, reasoning ability was found to be significant (F(1,86) =45.074, p= .000, power= 1.0). Secondly, visuo-spatial working memory was found to be significant (F(1,86) =35.654, p= .000, power= 1.0).

4.3.4 *Motor abilities*

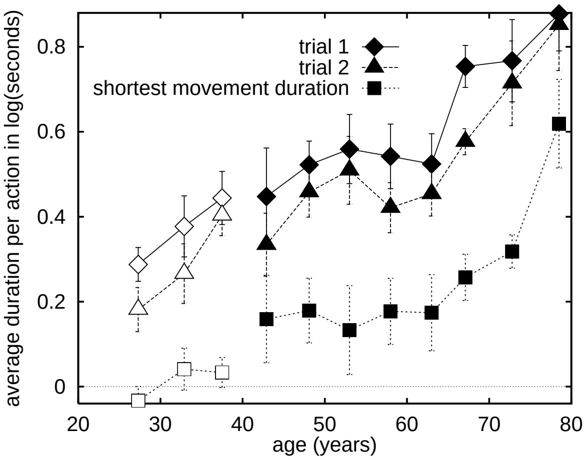


Figure 4.9: Movement duration measures per age group (log transformed).

In this study also motor abilities have been measured: Fitts' motion pretest, and the average duration per action within each session of the mobile phone task. Figure 4.9 depicts the transformed average duration per action per experimental trial. Furthermore, the transformed shortest movement duration per action achieved by subjects while clicking on squares is depicted in the same figure. Note that the curves of the average duration per action show a similar tendency as the shortest movement duration. All age groups took much more time to move from one button to another in the mobile phone experiment than in the Fitts experiment, i.e. than in moving from one square to another in the motion pretest.

4.4 Discussion

The results of this study indicated a technology generation effect, that is the discontinuity around the birth cohort of 1960, in the average number of mode errors, particularly in the first trial, and a continuous age effect in task duration while handling a three-layered user interface. This is in line with prior results (chapter 3) and the observed trends in the studies of Czaja [1996] and Czaja and Sharit [1998].

4.4.1 *Average error performance and task duration*

The transformed number of averaged mode errors is related to visuo-spatial working memory. Czaja and Sharit [1998] found a similar relation for procedural error performance. In the current study, the error performance within the software generation related also to the user's reasoning ability. However, the effects of age on visuo-spatial working memory found in our experiment does not correspond with those depicted in the literature overview of Salthouse [1991]. x

The factor visuo-spatial working memory in the study of Czaja and Sharit [1998] showed a similar continuous declining trend with age as that found in the overview of Salthouse [1991]. However, Czaja and Sharit composed this factor out of seven ability tests. Looking at Scheffé's test outcomes of these individual tests, four out of six tests that showed an age-related trend show a discontinuous effect. The Scheffé tests on Digit Symbol Substitution [Wechsler, 1981], the Trail Making Test form A [Reitan, 1980], Grooved Pegboard Test [Tiffin, 1968] and the Visual reproduction (delayed) [Wechsler, 1987] indicate that the youngest group performs better than the middle and oldest groups, but the middle and oldest groups do not differ significantly from each other.

This result raises the question why there has not been found any discontinuous effect in the overview of Salthouse [1991]. The figures in Salthouse [1991] are based on studies done before 1985 (e.g. [Wechsler, 1955; Wechsler, 1981; Schaie, 1985]. This period is just before the software style has been introduced widely to the market, even in the United States of America. Although the United States was a pioneer when it concerns computers, the first computer with a graphical layer and used by non-professionals, the Macintosh, has been put on the market only in 1984. Hence, the figures in the overview of Salthouse [1991] represent the abilities of the electro-mechanical generation. Also the figures of Schaie [1983] show that a 25 year old person in 1956 scores similarly well in spatial ability as a 25 year old in 1984. The results in our study and those of Czaja and Sharit [1998] may indicate a generational change after this period (1984) in the user's development of his/her visuo-spatial working memory.

Development studies [Sroufe and Cooper, 1988] already indicated that the adolescence and young adulthood are important periods in which experiences and cognitive advances influence each other intensively for further maturation. Having this framework in mind, it could very well be that the user interface developments and the increasing supply of complex devices in the formative period have given the software generation the opportunity to develop their visuo-spatial working memory much further than the electro-mechanical generation was able to do.

Which software-style features trigger a better visuo-spatial working memory? The visuo-spatial working memory test as executed in this study could give some suggestions as it was implemented in such a way that it would relate more closely to visuo-spatial working memory difficulties in software style devices. Firstly, it seems to be related to the need to memorize the visuo-spatial positioning of objects or information on a display. Secondly, also the visual distance between the mouse-pointing device and the system itself may play a role, inducing a more complicated hand-eye coordination, as part of parallel activities have to be done outside the user's visual field. The substantially shorter movement duration scores of the software generation compared to the electro-mechanical generation seems to confirm this generational advantage.

Our results on visuo-spatial working memory obtained by the computerized version gave the same effects as the similar paper-and pencil tests used by Czaja and Sharit [1998]. Consequently, the significant correlations that we found were not due to artificial similarity between the computerized Corsi test and the computerized mobile phone test. The difference between the computerized Corsi and Fitts scores shows that there are more effects in the Corsi than just the visual hand-eye coordination of the mouse. Hence, the complicated hand-eye coordination needed in our mobile phone experiment is just another factor that puts a heavy load on user's visuo-spatial working memory.

The average log-transformed task duration is related to both reasoning ability and visuo-spatial working memory. These findings are in line with the outcomes of Freudenthal, T.D. [1998] who observed that latency measures were related to visuo-spatial working memory and reasoning ability.

Speed-accuracy tradeoff

Results indicate that everyone learns to reduce the error rate and task duration. The significant interaction between age and trial for the transformed data suggests that something is going on with the user's learning ability. Analysis of the user's speed-accuracy tradeoff may give us a better indication of its developments. For this, the transformed mode error performance per trial has to be analysed in more detail.

Let us take a look at the errors in the first trial of the electro-mechanical generation (figure 4.9). It started with a relatively high error performance that decreased somewhat with age. Hence, older subjects tend to perform relatively more accurately than their younger generation members. This trend was not found in our previous study (see chapter 3), but can be observed in the results of Czaja [1996] and Czaja and Sharit [1998]. It also resembles the results of Welford [1968]. He observed that older people perform relatively more accurate as part of their speed-accuracy tradeoff. They achieve this by slowing down their speed of action (reflective strategy).

However, in the second trial the transformed error performance increased with age. This does not correspond with Welford's findings anymore. The converging trend within the electro-mechanical generation indicates that the oldest subjects were not able to learn that much from their mistakes despite their relatively more accurate performance at the start, compared to younger generation members.

The relation between speed and accuracy as found in this study is depicted per trial in figure 4.10. The average log-transformed duration per action is plotted on the x-axis and the transformed number of mode errors on the y-axis.

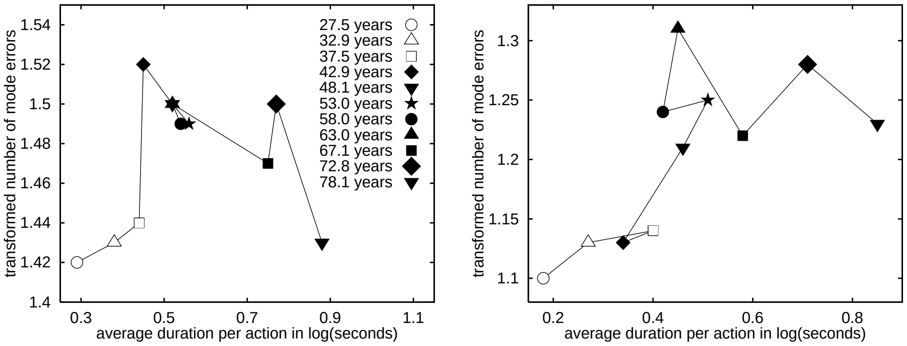


Figure 4.10: Relation between the user's average log-transformed duration per action and the transformed number of mode errors. Left-hand panel: First trial. Right-hand panel: Second trial. Note the different scales.

For a better understanding of the two pictures, imagine the following hypothetical speed-accuracy tradeoffs. A short duration combined with few errors is related to expert performance. A large number of errors combined with a short duration per action indicates a tradeoff of novice users who prefer speed above accuracy (speedy trial-and-error strategy). Fewer errors combined with a large duration per action suggest a tradeoff of novice users who prefer accuracy above speed (reflective strategy).

In the first trial, a moderate but insignificant increase with age of errors can be observed together with a substantial increase of duration per action until the average age of 43. From this age on a speed-accuracy tradeoff is visible: next to a large number of mode errors that decrease with age a steep increase of the averaged log-transformed duration per action is visible. Hence, in the first trial the trend within the electro-mechanical generation shows an age-related change from a speedy trial-and-error strategy to a reflective strategy. However, in the second trial the effect is different: next to an increase with age in the average duration per action, a more or less stable transformed number of errors can be observed for the first four age groups and another plateau of a higher transformed number of errors for the age groups above the age of 43.

Hence, a speedy trial-and-error strategy seems to have given the younger members of the electro-mechanical generation much more opportunity to develop an appropriate mental model of the system than the reflective strategy of the older users. This suggests that trial-and-error performance does not seem to be that bad after all for the electro-mechanical generation. It would also reduce the need to memorize visuo-spatial information during longer periods.

Moreover, the speed-accuracy tradeoff as found in this study confirms why the software generation is different from the electro-mechanical generation. The age groups belonging to the software generation have a short average duration per action combined with relatively few mode errors. Such speed-accuracy tradeoff is closely related to that of expert performance. This finding indicates once again that the software generation must have built up expertise and ability that gives them a priori advantage compared to the electro-mechanical generation when using a new device provided with a software style. This expertise may very well have a skilled visuo-spatial working memory and hand-eye coordination as components.

4.5 Conclusion

Visuo-cognitive skills seem to play a role in task duration. The transformed number of errors seems to be influenced particularly by technology generation-related differences in visuo-spatial working memory. Consequently, the software generation is able to act with a typical expert strategy. During the first confrontation with new technology, the youngest members of the electro-mechanical generation tend to use a speedy trial-and-error strategy, whereas the older members tend to use a reflective strategy. For the electro-mechanical generation the first strategy leads to a relatively larger learning effect.

4.5.1 *Implications for design*

The results of this study gave us an indication of the factors that have an impact on the usage of complex user interfaces. Although we cannot model these factors, as we do not know the exact weight of each factor, the outcomes of the analyses surely can give us ideas about the implications for design.

Designers need to be aware of the generation-specific ability gap between potential users born before 1960 and those born later concerning the user’s visuo-spatial working memory and hand-eye coordination. User problems of the electro-mechanical generation mainly, could be reduced by eliminating typical software style features that put a heavy load on user’s visuo-spatial working memory. Layers, long procedures, multi-functional buttons, and mode changes provoke inconsistent spatial placing of functions. Diverse (wireless) input and output devices such as the wireless remote control and mouse-pointing device seem to hamper a smooth hand-eye coordination, putting even more load on the user’s visuo-spatial working memory. Re-introduction of electro-mechanical style concepts such as the one-to-one relation between function and button, understandable direct feedback about the state of the device and increase of affordance could possibly reduce the load on visuo-spatial working memory.

4.5.2 *Implications for training*

A structured generation-specific (online) training could reduce the reasoning problems that the electro-mechanical generation face while learning to use software style devices. Focus of the training should be on a stepwise explanation of typical software style concepts and features to activate basic and specialized functionality. Furthermore, it is recommended to teach older users to explore instead of being cautious. Speeding up their action rate could reduce the time that they need to put a heavy load on their visuo-spatial working memory. This will facilitate them to develop a mental model that fits with the system’s conceptual model. However, how this learning strategy can be induced, for example by machine paced guidance, needs further study.

4.5.3 *Implications for further research*

This study indicated that the initial user strategy differs very much per age group. Not only between generations there is a different strategy, but also within the electro-mechanical generation. This suggests that the Welford’s concept of speed-accuracy tradeoff has lost its relevance in the era of software-based technology. Instead the concept of expert versus novice strategy has gained relevance. Analysis of the user’s approach to tackle an unknown device in terms of the user’s (erroneous) choices would help to understand the user’s mental model about software style devices. Therefore, in the next chapter the focus will be on the user’s action patterns using an even more complex software style device.

5

User strategies of the electro-mechanical generation

User strategies across four age groups of the electro-mechanical generation (born before 1960) have been studied while controlling a simulation of a new menu-driven television. The strategy of the menu generation (born in 1970 or later) has been taken as the norm, as these subjects grew up with menu-style devices. Within the electro-mechanical generation, the ineffectiveness increases with age (passive adaptation strategy). Those within the electro-mechanical generation who were effective (adapt actively) used an inefficient strategy that changed with age from a speedy trial-and-error strategy for the younger members towards a reflective strategy for the older ones. Also the learning strategy to tackle typical menu-style bottlenecks changed with increasing age, from successful navigation to a less useful manipulation. However, with practice also the older members of the electro-mechanical generation succeeded in changing their learning strategy from manipulation towards navigation.

5.1 Introduction

Whenever cable TV operators announce that they will change the frequencies of channels, it causes commotion among television viewers. Many of them, and in particular older users, have trouble with reprogramming channels on their television and video recorder and rearranging them in the order they were used to. Often children and grandchildren are asked for help. Those who do not have relatives, neighbors, or friends that can assist, may be forced to rely on paid services.

Two issues may be addressed in relation to this example. First, in real life, user interfaces of the television and video recorder are too complex for many television viewers [Freudenthal, A., 1998]. On present-day menu-driven user interfaces, users need to spend much effort and execute many actions to reach the promised functions. Secondly, young persons and older adults may differ substantially in their understanding of how current user interfaces and their programming procedures work, suggesting that user strategies of older adults may be less effective and less efficient than those of younger ones. Here, user strategy is defined as the user's approach to reach a predefined goal by means of the user interface of a certain device.

In prior research, reported in chapter 4, evidence has been found for generation-specific differences in user strategies. The results indicated that persons born in 1960 or later have an a priori advantage in handling current software style interfaces compared to adults born before 1960 due to the socio-technological environment they grew up with. In comparison with the older generation, the younger generation is more accurate and faster in using new appliances. This younger generation seems better equiped to deal with new devices provided with a software style interface as it grew up with such user interfaces during the formative years (i.e. between 10 and 25 years). This is the reason why Docampo Rama and Van der Kaaden [2001] identified those born in 1960 or later as belonging to the software generation. Those born before 1960 grew up with electro-mechanical interfaces, and were therefore identified as belonging to the electro-mechanical generation. In addition to differences between the two generations, we also found indications within the electro-mechanical generation for an age-related speed-accuracy tradeoff.

These studies concerned experiments of people using a simulated user interface featuring a simplified version of a present-day user interface (i.e. a less complex structure and less functionality than available on the market). Consumer electronics products on the market are more complex than those simulations. This raises the question whether the observed strategy differences also apply to menu-driven devices with a similar complexity as found in realistic daily life settings. This motivated the research described in this paper. Two main differences with the previous studies are the complexity of the user interface, caused by the increase in number of layers, and an extension of measures of user strategies to obtain a better insight in possible differences in user strategies. So what measures are relevant for the present study? The next section describes these measures.

5.1.1 *User strategies in daily life*

Effectiveness

According to ISO [1993] effectiveness is defined as the accuracy and completeness with which users achieve specified goals. It is often studied by focussing on the user intentions before dealing with a complex activity that determines whether a potential user will tackle the complex situation at all (e.g. Baltes and Baltes [1990]). Ineffectiveness occurs if either the user fails to start with a task or quits before the aim has been achieved. This latter aspect is referred in the literature as the user's coping attitude while handling the problem [Wister, 1989; Slangen-de Kort, 1999].

According to the literature, the effectiveness tends to decline with increasing age [Baltes and Baltes, 1990]. This has been explained in terms of the user's perceived cost-benefit change when confronted with complex situations [Baltes and Baltes, 1990; Melenhorst, 2001] and change in the user's adaptation strategy for situations in which environmental demands exceed the person's resources [Lawton, 1975; Lazarus and Folkman, 1984; Brandtstädter and Renner, 1990; Slangen-de Kort, 1999]. An active adaptation strategy involves active modification of the environment in the service of on-going goal attainment. This results in search for alternative ways to achieve the original goal. A passive adaptation strategy concerns accommodations of goals to losses and obstacles and results in giving up before achieving the goal. Wister [1989] found that with increasing age modes of adaptation shift from active to passive, increasing the probability of dropping out before having dealt with a given situation.

In this study, the effectiveness of users will be analysed by the number of subjects that persisted and succeeded while dealing with the menu-style system. Both failure and dropping out will be considered to be indicative of a passive strategy. Consequently, participants that show to be effective will be identified to have an active strategy. In line with the findings of Wister [1989], we expect the electro-mechanical generation to show an age-related increase in failure and drop outs, indicating an increasing probability in using a passive adaptation strategy as a function of age.

Efficiency

The user's strategy in terms of efficiency concerns the resources expended in relation to the accuracy and completeness of goals achieved [ISO, 1993]. In this chapter only the efficiency of effective users those who completed the task will be measured. The ISO-standard definition on efficiency only incorporates time-related measures. Here efficiency will be extended by taking the number of redundant actions and errors into account in line with Welford [1958] and Charness [1981].

In our previous studies, using a relatively simple user interface, the software generation has been shown to be both faster and more accurate. This reflects a level of efficiency that is typical for experts [Charness, 1981]. In contrast, the electro-mechanical generation as a whole showed an inefficient strategy by committing more errors and being on average slower than the software generation. In addition, within the electro-mechanical generation we found an age-related speed-accuracy tradeoff [Welford, 1958; Welford, 1968] in the sense that a speedy trial-and-error strategy changed with increasing age into a reflective strategy characterized by a slow but more accurate performance. Further practice resulted in a learning effect for all age groups, as both fewer errors and increased speed in which no speed-accuracy tradeoff was found.

In this study, it will be analysed whether a realistically complex system induces similar strategy differences within the electro-mechanical generation. The efficiency will be analysed by measuring the user's action duration and number of redundant actions. Experts are characterized by an efficient user strategy in terms of a short action duration and relatively few redundant actions. The electro-mechanical generation is expected to start with an inefficient user strategy until further practice. In line with prior research, the youngest groups of the electro-mechanical generation are expected to use a speedy trial-and-error strategy (large number of redundant actions in combination with a short action duration), whereas the older age groups will be expected to start with a reflective strategy (long action duration in combination with relatively few redundant actions). All age groups are expected to operate more efficiently (fewer redundant actions, shorter action duration) with further practice, irrespective of their initial user strategy.

Bottleneck handling

A third aspect of user strategies concerns the approach to overcome difficulties. This is referred in this study as *bottleneck handling*, and can be analysed by measuring how many and which action choices have been made to overcome bottlenecks while tackling a predefined goal [Baxter and Oatley, 1991].

In a study by Westerink, Majoor and Docampo Rama [2000] the action choices of inexperienced CD-i users and the way the choices change with practice were studied in terms of menu paths chosen (navigation) and content viewed (manipulation). For this, subjects between 20 and 44 years could explore first and we asked to search for specific information in an encyclopaedia on CD-i that was organised according to a menu structure. Users spent most of their time looking for the composition of the menu by navigating through many paths without manipulating options to view information. This occurred both while searching for a priori specified information as well as during the exploration phase. Only after understanding the structure of the menu, the exploration strategy changed into explicit viewing and manipulation of the content. Hence, the user needs to understand the menu structure for confident use of functionality. Consequently, navigating instead of manipulating appears to be the suitable learning strategy to overcome typical menu-style difficulties (bottlenecks).

In the present study, the goal was to assess the percentage of redundant navigations (relative to the overall number of redundant actions) as an indicator of the user's efficiency in tackling menu-style difficulties. In addition, the user's actions will be analysed in more detail by identifying what steps induced a substantial amount of redundant action (these are then the most prominent bottlenecks), and which types of redundant actions (navigations, manipulations, and mode errors, see figure 5.4) were executed while searching for the solution.

In this study, the percentage of redundant navigations executed within the electro-mechanical generation is expected to decrease with increasing age; the youngest groups of the electro-mechanical generation are expected to overcome obstacles similar to the menu generation, i.e., learning by a high percentage of redundant navigations. On the contrary, older age groups are expected to show a low percentage of navigations while reflecting about the key button to press that activates (manipulates) the function in mind directly.

5.1.2 *Focus*

The present study concentrates on age-dependent user strategies within the electro-mechanical generation while handling a realistically complex menu-driven user interface. A menu-driven television will be simulated. The setting will reflect a situation in which the user cannot rely on help from outside.

The user strategy of the menu generation (born in 1970 and later) will be taken as a reference. Their members grew up with menu-driven user interfaces and will therefore show high effectiveness, high efficiency and few bottlenecks with an efficient navigational learning strategy.

Within the electro-mechanical generation the user strategy will be analysed by the age group's (1) effectiveness, measured per age group by the number of persisting participants and number of subjects who successfully completed a (sub)-task (part of ISO standards), (2) efficiency of those who succeeded in terms of the average action duration (according to ISO standards), and in terms of number of redundant actions executed, and (3) bottleneck handling, which will be measured in terms of the percentage redundant navigations executed to fulfill the task, and identification of the steps in the sequence that particularly induced a large number of redundant actions. These three measures will indicate whether an active or passive strategy is used, whether inefficiencies are based upon a speedy trial-and error handling or a reflective strategy, and whether typical menu-style bottlenecks are tackled using an efficient navigational learning strategy.

5.2 Method

5.2.1 *Subjects*

The experiment was executed in 1999. Fifty Dutch subjects between the ages of 20 and and 80 years took part and were assigned to five age groups of ten subjects each. The youngest age group of 20 to 25 years the reference group being member of the menu generation. The age groups of 45 to 50 (younger members of the electro-mechanical generation and part of the working force), 55 to 60 (still part of the working force), 70 to 75 (retired for some time), and 75 to 80 years (oldest age group) were all part of the electro-mechanical generation. Educational level was confined to high school diploma or bachelor degree (both equally present per age group), resulting in a homogeneous experimental population in this respect. An equal number of female and male subjects per age group was selected to control for possible gender differences. Subjects did not have any self-reported problems with manual movements, hearing, or vision.

Figure 5.1 indicates per age group two visuo-cognitive abilities of the subjects that may be relevant for their task. The reasoning ability for visual material (left-hand panel) was assessed by Raven's Progressive matrices test [Raven, 1960]; the user's visuo-spatial working memory (right-hand panel) was measured with a computerized version of Corsi's block tapping test [Milner, 1971], see chapter 4 for details. As can be seen in the left-hand panel of figure 5.1, the Raven's score for visual reasoning ability declined systematically with age (F(1,48)= 19.93, p= .000). Compared to normative studies like that of Burke [1985], these subjects scored above average, which is probably related to their high educational level [Vernon, 1947]. Visuo-spatial working memory (right-hand panel of figure 5.1) differed significantly across the two generations with the youngest generation scoring highest on it (F(1,48)= 10.04, p= .002). Within the electro-mechanical generation, the visuo-spatial working memory scores did not seem to decrease with age. Hence, in general, the menu generation scored higher on these visuo-cognitive ability tests. The visuo-cognitive scores in the present study resemble the ones found in a prior study described in this thesis, in which different subjects with similar educational background participated.

An explicit difference between the menu and electro-mechanical generations is their experience in handling menu-driven devices. As a menu-driven television was used in this study it was asked whether the subjects had used the menu of a television before. Seven out of ten subjects of the menu generation reported some experience handling menu's on a television. Within the electro-mechanical generation exactly three subjects for each age group had used the menu of a television before.

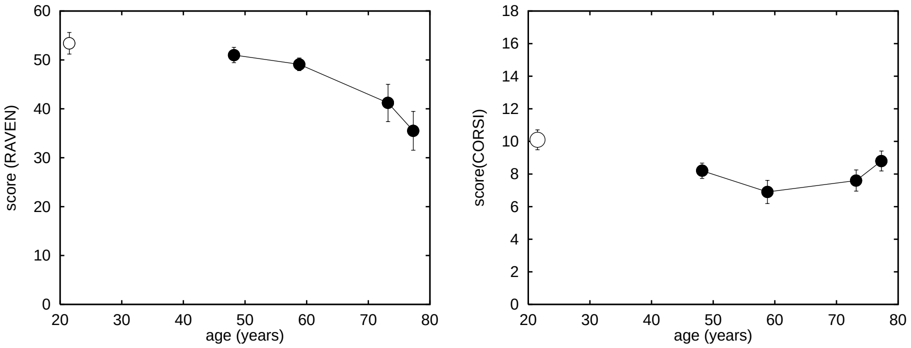


Figure 5.1: Visuo-spatial ability scores depicted per age group. On this and the following figures, open symbols refer to the menu generation, black symbols refer to the electro-mechanical generation, and vertical bars denote twice the standard error of the mean. Left-hand panel: Average score on reasoning ability, measured by the Raven's complex matrices test [Raven, 1960]. Right-hand panel: Average score on visuo-spatial working memory, measured by a computerized and modified version of Corsi's block-tapping test [Milner, 1971]. For details on modification, see chapter 4 of this thesis. Each point in these and following figures indicates the average scores of ten subjects.

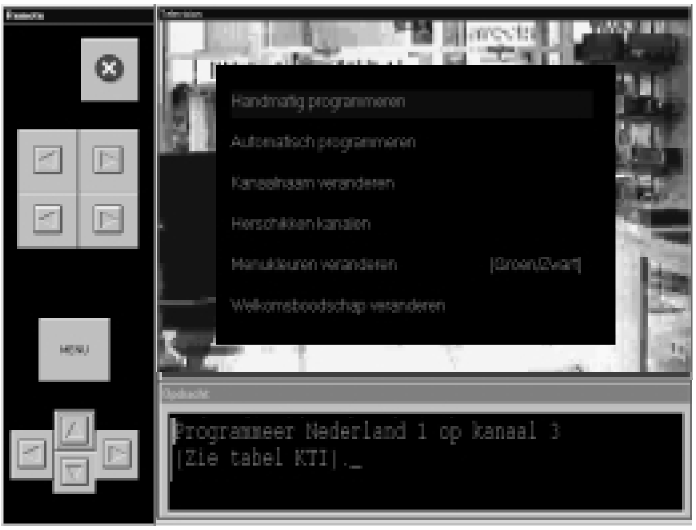
5.2.2 *Materials*

Three menu-driven televisions with remote control (RC), that just had appeared on the market, served as examples for creating a user interface simulation with a realistic complexity level of four layers, including the upper layer visible on the RC (layer 0) and three menu layers visible on the television screen. The simulation was built in Microsoft Visual Basic 5.0 in the Windows 95 environment. Figure 5.2 depicts the simulation that is presented full-screen on a 17 inch 1024 x 768 pixel color monitor. The subjects were placed at a distance of 50 cm from the monitor. Although a real remote control would have been more appropriate for the setting, only a mouse-pointing device for activation of the RC could be realised for this study. The simulation is composed of three windows boarding each other. The left window represents the RC. The upper right window is the tv screen on which a menu can be visualized over a video sequence that is linked to a channel. In the window below, the textual description of the task is presented to the subject.

The simulated RC has 10 buttons that have a minimum size of two by two cm, and can be activated by clicking with a mouse-pointing device without precision. The buttons represent from top to bottom: power on / off, channel up and down, volume up and down, the menu, and arrow up, left, right, and down. They serve three basic functions, namely tv-power change, channel change, and volume change, and six specialized functions: manual programming, automated programming, reorder channels, channel name change, menu colour change, and welcome message change.

The basic functions can be manipulated directly. All specialized functions can only be manipulated on the RC by navigating through a menu that can be visualized on the television screen via the menu button.

The menu proper of the simulated television consists of three layers each with words with a font size of 20 points (1 cm). On the first layer of the menu six menu options are visualized referring to the six specialized functions. This menu is depicted in figure 5.2. Automated programming and change of menu colours can be manipulated on the second layer of the menu. All other options can only be manipulated on the third layer of the menu. Button presses that do not activate a function in a certain mode result in mode errors. These can be committed when the arrow buttons are pressed without activating the menu, or when the channel



Programmeer Nederland 1 op kanaal 3
(Zie tabel KTI)._

Figure 5.2: Black-white version of the user interface of the television simulation. The buttons on the RC represent from top to bottom: power on/off, channel up/down, volume up/down, menu on/off, arrow up/down/left/right.

and volume buttons are pressed after activating the menu. The television can be programmed for a maximum of 99 channels.

A logging tool registered each button press executed by the subject and the time between button presses. The experiment itself was run on a Pentium MMX 200. The subjects were seated in a quiet room that was well lit without specular light reflection on the computer screen.

5.2.3 *Description of simulation in terms of menu-style actions*

Figure 5.3 depicts the organization (task model) of the function 'manually programming a channel'. The Méthode Analytique de Description (MAD) [Scapin and Pierret-Golbreich, 1990; Sebillotte and Scapin, 1994; Sebillotte, 1995] is used to graphically represent the organization of such function. The lowest level of the task model represents the user's button action sequence on the remote control (RC) to reach the goal (manually programming a channel). Higher levels of the task model represent the underlying structure behind the user actions. This structure is in fact what the user should understand to be able to use the menu style interface without problems.

On a menu style interface, specialized functionality like programming a channel is organized in a standard format using navigational buttons (menu button, arrow up, down, left, and/or right) according to a sequential procedure that is composed of a navigation and a manipulation component. The organization of a menu style interface in its totality (as described in the historical analysis of Docampo Rama and Van der Kaaden [2001]) visualized by MAD is depicted in figure 5.4.

Figure 5.4 shows the menu style divided in three types of functions: no function at all (mode errors), basic functions, and specialized functions. Basic functions concern only manipulations (power off / on, volume change, and channel change). Specialized functions contain a sequence of navigational actions (selections that lead to a deeper layer in the menu, option shifts to go through the option list, and/or backward navigations that result in going up one menu level) and manipulations (channel setting, channel switching, menu color change, or character changes).

The dotted rectangle at the right in figure 5.4 represents the task model of specialized functionality composed of a sequential, layered procedure. Obviously, the procedure for specialized functionality is more complex than that of basic functionality.

5.2.4 *Procedure*

The subjects had to execute three tasks. Task 1 consisted of 3 subtasks that had to be repeated once in the same order. Task 2 and 3 consisted both of 2 subtasks

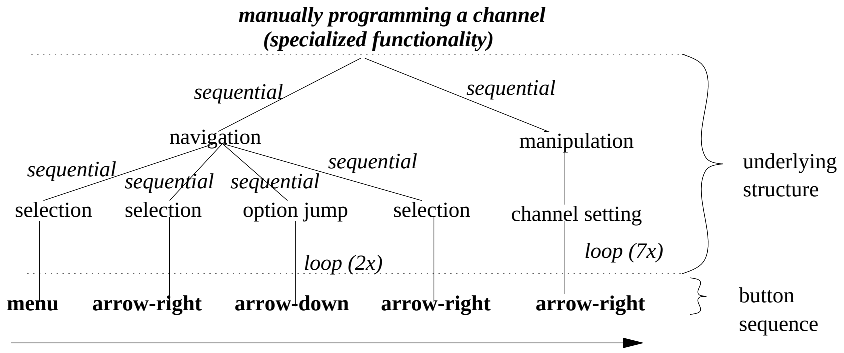
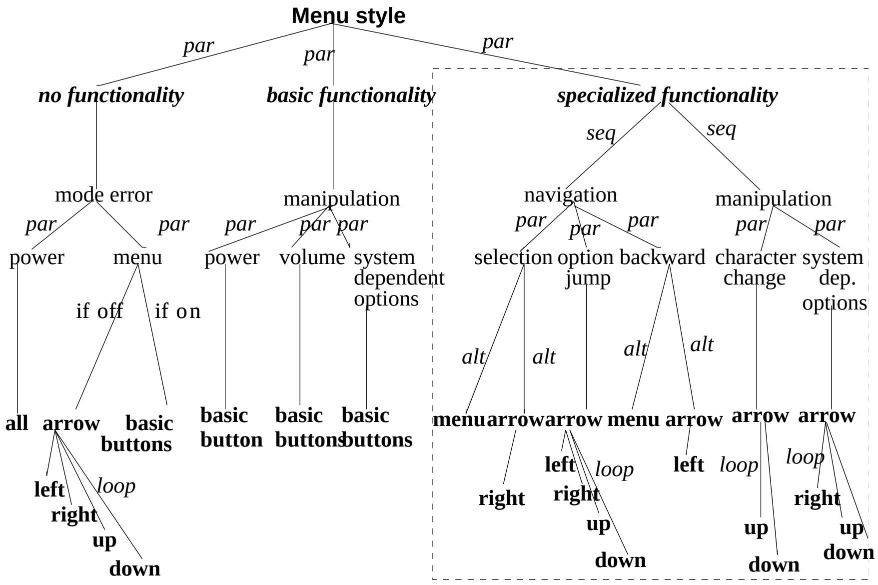


Figure 5.3: Example of a task model corresponding to the specialized function 'manually programming a channel' on a menu-driven television like the one that will be used in this study. It is the first subtask of the experiment. The button sequence represents the order of actions from left to right. The word 'sequential' means that the action has to be executed according to the order as it is indicated in the figure (meaning that you need to go down in the tree to know which actions have to be carried out before going on with the following one). So, first the menu has to be activated, then the right arrow, etc. After the first 'arrow-right' the user will have selected the option to program. After the second 'arrow-right', the user will have selected the channel that is asked for. The terms on the level of the button sequence (like arrow-right) correspond with the buttons as described in figure 5.2. For other specialized tasks the button sequence will be different, but the underlying structure remains similar, namely starting with navigations en finishing with manipulations.

that had to be carried out twice in a fixed order as well.

The order of the subtasks within tasks 1, 2, and 3 resembled a realistic setting in daily life: a new television, of which the cables are already plugged in, and the settings have to be programmed first before being ready for normal use. Therefore, task 1 of the experiment consisted of 'Programming channels' (minimally 4 sub-procedures, 18 button presses), then 'Rearranging the channels' (minimally 5 sub-procedures, 9 button presses), and finally 'Giving the channels a name' (minimally 6 sub-procedures, 81 button presses). Task 2 concerned two subtasks of specialized functionality that are found only on the latest televisions, which is 'Changing the colours of the menu' (minimally 3 sub-procedures, 5 button presses) and 'Programming a welcome message' (minimally 3 sub-procedures, 67 button presses). Task 3 consisted of two subtasks belonging to basic functionality, which are 'Watching specific tv channels' (minimally 1 sub-procedure, 16 button presses) and 'Changing specific volume levels' (minimally 1 sub-procedure, 19 button presses). To carry out task 3 the users only had to use channel up and



bold-italic font = generic functionality
normal font = interaction structure
bold font = interaction techniques (button presses)

italic font = types of procedures corresponding to node
par = parallel tasks (no fixed order)
seq = sequential tasks (fixed order)
alt = alternative or optional procedure for same task
loop = routine tasks (loop of same button presses to reach the goal)

Figure 5.4: Decomposition of the menu style of consumer electronics products as described in a historical analysis [Docampo and van der Kaaden, 2000]. Its graphical representation is build up according to MAD [Scapin and Pierret-Golbreich, 1990; Sebillotte and Scapin 1994; Sebillotte, 1995]. The rectangle at the right highlights the task model of specialized functionality. Explanation from left to right: the arrow-buttons have no function when the menu is off. Basic functions, except for the power button, have no function when the menu is on. Basic functions are the power button, volume, and system dependent functions like the channels for a television (see figure 5.2). Specialized functionality consists always of a sequence composed firstly of at least one navigation and secondly of at least one manipulation.

down, and volume up and down buttons. Hence, no menu had to be used for these final subtasks. This all is summarized in figure 5.5, showing all subtasks per task, corresponding optimal sub-procedures (sequence of steps), and minimal number of button presses (actions).

If subjects did not finish a subtask within 10 minutes, they were assisted in

tasks	subtasks	explanation procedure	sequence of steps	minimal actions
task 1: specialised	programming	1. select menu 2. select program option 3. go to channel for programming 4. program channel frequency	1. navigation (selection) 2. navigation (selection) 3. navigation (optionjump) 4. navigation (selection) manipulation (channel setting)	18
	switching	1. go one menu layer up 2. go to switch option 3. select switch option 4. go to channels for switch 5. switch channels	1. navigation (backward) 2. navigation (option jump) 3. navigation (selection) 4. navigation (option jump) 5. navigation (selection) navigation (option jump) manipulation (channelsetting)	9
	naming	1. go one menu layer up 2. go to naming option 3. select naming option 4. go to channel for naming 5. select channel 6. name channel	1. navigation (backward) 2. navigation (option jump) 3. navigation (selection) 4. navigation (option jump) 5. navigation (selection) 6. navigation (option jump) manipulation (characters)	81
task 2: specialised	menu colors	1. select menu 2. go to menu color-option 3. change colors of menu	1. navigation (selection) 2. navigation (option jump) 3. manipulation (menu colors)	5
	message	1. go to message option 2. select message option 3. change message	1. navigation (option jump) 2. navigation (selection) 3. navigation (option jump) manipulation (characters)	67
task 3: basic	channels	1. watch program x, y, z	1. manipulation (channel)	16
	volume	1. change the volume	1. manipulation (volume)	19

Figure 5.5: Task decomposition in subtasks and their minimal sequence of steps (sub-procedures) expressed in categories of menu-style actions (button presses). The last column indicates the minimum number of actions required to fulfil the task.

starting-up the next subtask. A pilot experiment [Schram, 1998] showed that subjects should not be kept searching any longer because they are unlikely to reach the goal and quit anyhow. The subjects who failed were always encouraged to continue with these follow-up tasks.

Before starting the experimental trials, all subjects were given a general instruction and exercised with the mouse-pointing device. The entire experiment took on average one hour and thirty minutes.

5.2.5 *Data transformation towards user strategies*

A logging tool, attached to the simulation, registered the buttons pressed by the user and the time interval between two button presses (action duration). As seen in figure 5.4, this is a low-level description of the user's executed actions. As buttons on a menu-style interface can have several functions, depending on their mode, such low-level descriptions do not tell what types of actions were achieved by the user after pressing certain buttons. For this, the data needs to be further transformed into high-level descriptions of these sub-procedures.

Two automated tools have been used after the experiment for the transformation. These tools concerned a conversion tool and a counting tool. Figure 5.6 depicts a simplified explanation of the contribution of each tool during the process of data transformation.

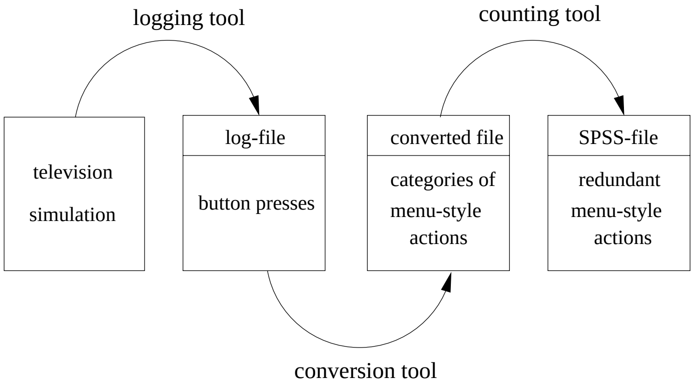


Figure 5.6: Procedure for the analysis of user strategies by running three fully automated tools. The button presses were converted per step and trial. The figure does not show the action durations between two button presses. These were logged as well, but not converted per step. In the SPSS-file action durations are registered only as average per task and trial. Just to give an example of the registration in each part of the process, think of subtask 1.1 described in the caption of figure 5.2. If a user presses the power button off and on again before fulfilling the sequence as described in that figure, the log file will log these three actions, plus menu-on, arrow right, arrow-down, and arrow-right (eight times). Moreover, the action duration between each action is then registered as well, next to their total task duration. All this will then be converted serially into manipulations - power (two times), navigation - selection (three times) and manipulations - channel settings (seven times), plus the total amount of time. Then the redundant steps are automatically counted and entered in a SPSS-file: number of redundant actions in trial 1 for task 1.1: two; redundant manipulations: two; redundant power: two; all other types of actions: zero. Also the total number of actions executed by the user and the total amount of time for that task are entered in the SPSS-file.

Conversion tool: As can be observed in figure 5.4, one type of button press can have several functions depending on its mode. Therefore, a conversion tool has been developed based on Layered Protocol Modelling [Taylor, 1988a; Taylor, 1988b; Taylor, 1989; Haakma, 1998; Brinkman, 2000]. This method consists of an interaction component that rebuilds relevant parts of the conceptual model of the system from the actual interaction with the subjects. That is, this component translates button presses unambiguously into the user's executed categories of menu-style actions as identified in figure 5.4, which are navigations, manipulations and mode errors. Navigations are subdivided in selections, option shifts and backward navigations. Manipulations can be classified in this study as power change, volume change, channel change, channel setting changes, character changes, and menu colour changes. Note that action durations have not been converted.

Counting tool: In this study we will capture user strategies while handling the system by means of a counting tool. It scores the total number of redundant actions per menu-style category (navigations, manipulations, mode errors), which are registered per step, subtask and task to be reached.

5.3 Results

5.3.1 *Effectiveness*

Table 5.1 depicts per age group the number of subjects that started (shown between brackets) and succeeded per task and trial. None of the subjects refused to start with the first subtask. From the 40 subjects of the electro-mechanical generation, mainly the oldest age group of 75-80 dropped out before finalizing the experiment: 3 out of 10 subjects of this oldest group stopped immediately after reaching the first step (which is the activation of the menu), and 4 others did not continue after finishing the first trial of the first task (programming, switching channels, and giving a name) when they heard that they had to repeat the exercise (table 5.2). Except for the persons in the age group of 70-75, most other subjects within the electro-mechanical generation stayed on, just like the whole menu generation.

But how many subjects indeed achieved the goals? Except for the first trial of task 1, all members of the menu generation have been found to be effective in this respect. The electro-mechanical generation failed more often compared to the menu generation. During the first task, the number of subjects that failed within the electro-mechanical generation increased with age. But as table 5.1 shows, those of the electro-mechanical generation who continued in follow-up tasks became more effective in subsequent tasks and particularly in the second trial of each task. For both the electro-mechanical generation and the menu generation trial 1 of task 1 was most difficult to overcome.

Table 5.2 depicts the effectiveness per trial of the subtasks of the most difficult task to achieve: task 1. Three trends in effectiveness can be seen within the electro-mechanical generation. First, the effectiveness seems to decrease with increasing age in terms of an increasing number of failing subjects and drop outs.

Table 5.1: Number of subjects across age groups that succeeded per trial and task.

age group	task 1		task 2		task 3		
	trial 1	trial 2	trial 1	trial 2	trial 1	trial 2	full
20-25 y.	8 (10)	10 (10)	10 (10)	10 (10)	10 (10)	10 (10)	8 (10)
45-50 y.	8 (10)	10 (10)	10 (10)	10 (10)	10 (10)	10 (10)	8 (10)
55-60 y.	6 (10)	8 (10)	10 (10)	10 (10)	9 (10)	10 (10)	4 (10)
70-75 y.	1 (10)	6 (8)	7 (8)	8 (8)	6 (7)	7 (7)	0 (10)
75-80 y.	2 (10)	3 (3)	3 (3)	3 (3)	3 (3)	3 (3)	2 (10)

Number of subjects who started task in brackets, full = full completion of all tasks, task 1 and 2=specialized functionality, task 3=basic functionality,

Table 5.2: Number of subjects across age groups that succeeded per trial and subtask of task 1.

age group	subtasks of trial 1			subtasks of trial 2		
	program	switch	name	program	switch	name
20-25 y.	8 (10)	9 (10)	10 (10)	10 (10)	10 (10)	10 (10)
45-50 y.	8 (10)	10 (10)	10 (10)	10 (10)	10 (10)	10 (10)
55-60 y.	6 (10)	9 (10)	10 (10)	8 (10)	10 (10)	10 (10)
70-75 y.	5 (10)	3 (10)	5 (10)	7 (8)	6 (8)	7 (8)
75-80 y.	6 (10)	5 (7)	4 (7)	3 (3)	3 (3)	3 (3)

Number of subjects who started task in brackets, program= programming a channel, switch= switching channels, name= naming channel.

The data suggest that the youngest age group of the electro-mechanical generation (45-50 years) is similarly effective as the menu generation. In this study, the two oldest age groups (70-75 and 75-80 years) within the electro-mechanical generation seem to differ most from the menu generation. Second, most drop outs were seen at the start. Third, those of the electro-mechanical generation that persisted with follow-up tasks ended up nearly as effective as the menu generation, indicating that all effective members of the electro-mechanical generation learn.

A question is whether the persisting subjects (those who use an active strategy) of the older age groups show a different (higher) reasoning capacity than the drop outs (those who use a passive strategy). The Raven and Corsi scores of the persisting subjects of the two oldest age groups (respectively 42.5 and 9.5 on average) appear somewhat higher than the Raven and Corsi scores of the drop outs (respectively 38 and 8 on average), but the difference fails to reach significance due to the low number of persisting subjects (F(1,18)= 0.241, $p = .63$).

Table 5.3: Number of redundant actions needed per task averaged per generation.

generation	task 1	task 2	task 3
menu - with outlier	199	9	23
menu - without outlier	149	0	4
electro-mechanical - no outlier	417	30	120

5.3.2 *Efficiency*

The users' efficiency has been analysed per task and trial by measuring the average action duration and the number of redundant actions of the effective subjects.

Table 5.3 shows the average number of redundant actions in task 1, 2 and 3 (both trials summed). Hence, the electro-mechanical generation made substantially more redundant actions than the menu generation. Detailed analysis of the menu generation showed that the relatively large number of redundant actions in task 1 and during the first trial of task 3 were made by one particular subject. No such extreme outliers were found within the electro-mechanical generation.

The users' specific speed-accuracy tradeoff per task and trial can be identified by plotting the number of steps needed to the average duration per action. That is what figure 5.7 depicts per task and trial. The general trend is that the first trial of task 1 is performed inefficiently. After that, a drastic reduction of redundant actions and of action duration was found in the second trial and while achieving task 2. This suggests that, like the menu generation, also the electro-mechanical generation is able to learn by experience and to transfer the specialized procedures from task 1 to new functions with similar procedures in task 2. Again an increase in action duration and number of redundant actions is seen to activate basic functions in task 3. Apparently, a change of procedures from those for specialized functionality to those for basic functionality (concerning the transfer from task 2 to task 3) induces a relatively inefficient user strategy that is corrected in the second trial. Within the electro-mechanical generation there is a systematic increase with age in the average action duration, whereas the average number of redundant actions tends to increase with age only for the first two age groups of this generation, and declines somewhat for the oldest age groups. This resembles the results of prior research reported in chapter 3 and 4. And again it is observed that the younger age groups within the electro-mechanical generation prefer to start with short action durations and many redundant actions (speedy trial-and-error strategy), whereas the oldest age groups tend to be slower but perform relatively more accurately (reflective strategy).

As mentioned before, only three out of 10 subjects of the oldest age group persisted during all tasks. This suggests that these three subjects best understood

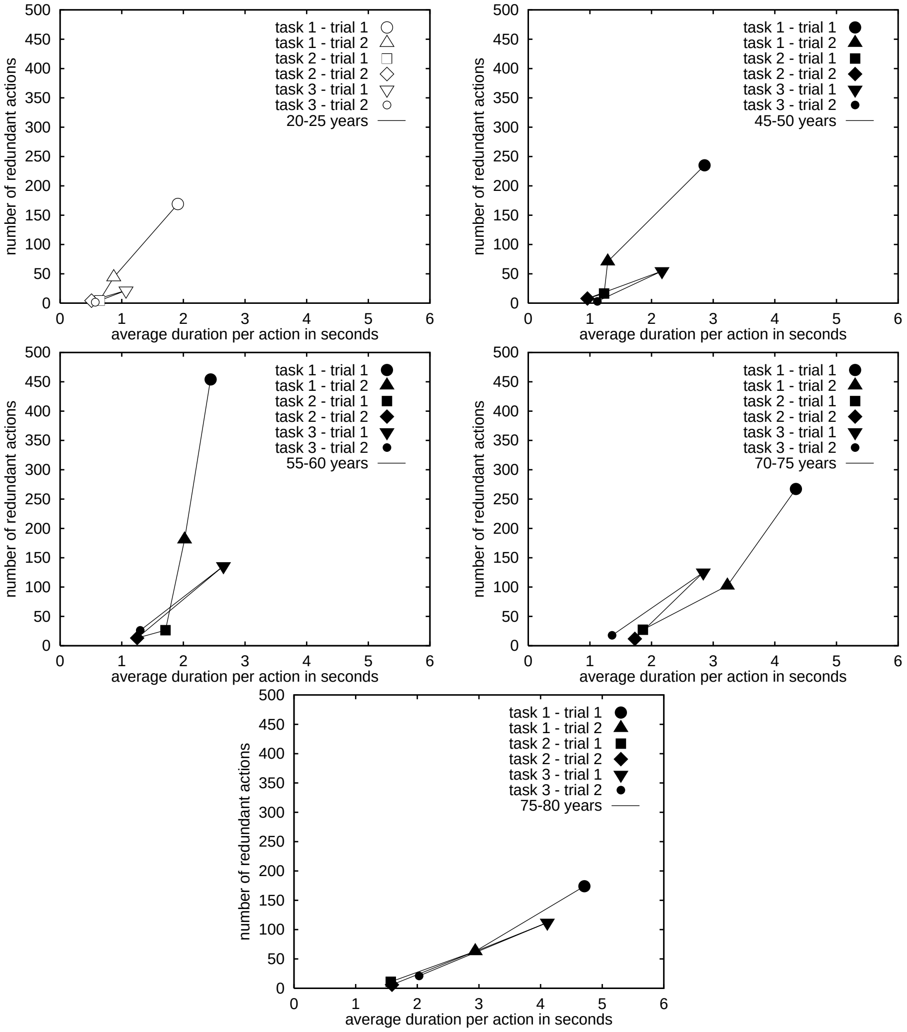


Figure 5.7: Task efficiency in terms of action duration and redundant actions.

the tasks, whereas other groups would consist of a mix of subjects performing more heterogeneously. One way to check this is to compare the performance of the 3 most accurate subjects across age groups. Such a comparison can give us an indication of the difference in performance between the most accurate subjects and the remaining subjects. Due to the low number of subjects, no statistically significant results can be expected. As can be seen in table 5.4, a similar pattern is found across tasks and trials as in figure 5.7: a decrease in redundant actions and action duration from task 1 to 2, and an increase from task 2 to 3. The differ-

Table 5.4: Number of redundant actions and action durations in seconds averaged for the best three subjects per age group.

measure	age group	task 1	task 2	task 2	task 3	task 3
trial		2	1	2	1	2
redundant actions	20-25 y.	10	2	2	0	0
	45-50 y.	13	4	2	3	0
	55-60 y.	30	11	3	17	0.7
	70-75 y.	49	9	2	39	0.7
	75-80 y.	63	11	6	112	21
action duration	20-25 y.	0.7	0.7	0.4	0.9	0.6
	45-50 y.	0.9	0.8	0.8	2.3	0.9
	55-60 y.	2.0	1.4	1.1	2.2	1.1
	70-75 y.	2.9	2.9	1.6	4.1	0.8
	75-80 y.	2.9	1.8	1.6	4.1	2.0

Redundant actions represent the total number of redundant actions for the second trial of task 1, and trial 1 and 2 of task 2 and 3. Task 1 trial 1 has not been incorporated, as the two oldest age groups failed to acquire at least 3 successful subjects at that stage of the experiment.

ence between the general trend and the performance of the most accurate subjects of each group is that both the number of redundant actions and action duration increase monotonically with age (accurate subjects), instead of discontinuously with age (general trend).

In summary, the effective subjects of the electro-mechanical generation show a less efficient strategy than the menu generation. Within the electro-mechanical generation there is a speed-accuracy tradeoff, i.e. a speedy and inaccurate performance of the younger groups and a more careful and more accurate performance by the older ones. All age groups show a clear-cut learning effect in follow-up trials and tasks with similar procedures. However, the follow-up task 3 with a different (even more simple) procedure is perceived by all age groups within the electro-mechanical generation as a bottleneck; it induced a persistence effect in terms of user strategy, resulting in less efficient behaviour (larger action duration and an increase in number of redundant actions). From the menu generation, only the outlier showed a similar performance.

5.3.3 *Bottleneck handling*

The results showed that task 1, which concerns the first series of specialized functionality, and task 3, which concerned basic functionality, have been found difficult to complete. Therefore, the first trials of the subtasks are analysed further,

Table 5.5: Average percentage of navigations of the total number of redundant actions in the first trial of each subtask of task 1 and task 3, depicted per age group.

age	progr. (t1)	switch(t1)	name (t1)	watch (t3)	volume (t3)
20-25	51 (8)	52 (9)	91 (10)	44 (10)	0 (10)
45-50	52 (8)	62 (10)	90 (10)	50 (10)	36 (10)
55-60	57 (6)	50 (9)	90 (10)	56 (10)	31 (10)
70-75	35 (5)	50 (3)	85 (5)	68 (7)	0 (7)
75-80	37 (6)	60 (5)	88 (4)	65 (3)	0 (3)

progr. = manually programming a channel, switch = switching channels, name = naming channels, watch = watching programs, volume = changing volume, t1 = task 1, t3 = task 3, numbers between brackets represent the number of subjects that participated in the particular subtask.

as most difficulties were spotted there. Task 1 concerned programming a channel, switching channels, and naming them. Task 3 concerned watching specific channels and changing the volume.

In task 1, the goal is to study what percentage of the redundant actions concerns navigational actions, as these are indicative of a useful learning strategy for typical menu style functions.

As Table 5.5 shows, in the very first subtask the two youngest age groups of the electro-mechanical generation showed a similar percentage of redundant navigations as the menu generation (slightly over 50 percent). The oldest two age groups produced relatively fewer redundant navigations and more redundant manipulations. However, in following subtasks within task 1, the two oldest age groups were able to recover. Only in the third subtask of task 1, all age groups executing nearly 100 percent redundant navigations. Clearly, the subjects had then learned that navigation is a useful strategy.

In contrast, no typical menu-style actions were needed in task 3 and redundant navigations were inappropriate. Although task 3 contains only basic functions, all age groups started treating it as a typical menu-style function learned and rewarded in all previous trials. The percentage of redundant navigations produced was the highest for the two oldest age groups (around 65 percent compared to 53 percent for the younger age groups of the same generation). However, during the second subtask of task 3, the oldest age groups were able to apply earlier their familiar basic learning strategy of manipulations only. Now the oldest age groups were able to perform similarly well as the menu generation.

The large number of redundant navigations during all subtasks may have been produced on purpose to find out the structure of the menu, but may also indicate the user's problem in understanding how this navigational procedure works. In the latter case, the user would press the navigation buttons without understanding what has happened. Certain reactions of older users during the experiment seem to suggest this latter pattern. Examples of such reactions were; ”what did I do now?”, ”why doesn’t it do what I want him to do?”, and ”I have no idea what these buttons are for”. What we observed was that subjects tended to replicate part of the previously learned procedures.

Table 5.6 gives for each age group an overview of the types of redundant actions (navigations, manipulations, mode errors) per step of the sequence of each subtask of task 1 (programming a channel, switching channels, and naming a channel). For the first subtask (programming a channel) the electro-mechanical generation perceived 3 main bottlenecks (in terms of a large number of redundant actions) during the first subtask of task 1: step 2, 3 and 4. With experience in follow-up subtasks the number of bottlenecks decreased, and appeared rather at the end of the sequence. Spontaneous comments of the subjects suggest that finding the procedure behind the manipulative action caused the trouble during the final step.

Note that the redundant actions executed to achieve step 2 of the first subtask (programming a channel) mainly consisted of manipulations. The menu generation used a typical menu-style manipulation, namely the automated program feature. However, the redundant manipulations within the electro-mechanical generation often concerned basic functionality like going from one channel to another, and manipulations for all types of menu-style functions, other than programming a channel.

The table shows only the difficulties per step for those subjects who, at the end, succeeded anyhow. However, real bottlenecks were those steps in which subjects were not able to succeed at all. During the first subtask ineffective subjects happened to fail most often while dealing with the final step of programming manually, but three subjects already failed dealing with step 1 (activating the menu itself). For switching channels, the ineffective subjects were unable to overcome the selection of the channels to switch and the switching act itself. Failure during the third subtask consisted mainly of being unable to change the characters to name the channel (which is the final step).

Table 5.7 depicts the types of redundant actions per step for the subtasks of task 3 (channel change and volume change). The redundant actions of the first subtask (watching a tv program, which is a basic function) consisted of both navigational actions and manipulations. Further analysis (not visualized in the table) shows that for most age groups except for the oldest age group, these redundant manipulations concerned mainly menu-based manipulations (channel setting changes). In contrast to the other groups, the redundant manipulations of the oldest group consisted mainly of power changes (switching the television on and off).

Hence, at the start, the learning strategy of the youngest age groups of the electro-mechanical generation was most similar to the menu generation. But the older age groups were able to recover. Change of procedure shows a persistence of the learned procedure for menu style functions. It is the oldest age group of the electro-mechanical generation that is able to adapt the basic procedure like the menu generation did.

5.4 Discussion

This chapter deals with age-dependent user strategies of the electro-mechanical generation when handling realistic user interface complexity. The user strategies have been analysed by the age group's effectiveness, efficiency, and bottleneck handling.

5.4.1 *Effectiveness*

Within the electro-mechanical generation the effectiveness drops with increasing age. This indicates that the adaptation strategy tends to change with increasing age from active to passive, in line with Wister [1989]. In fact, only the youngest age group of the electro-mechanical generation was able to adapt as actively as the menu generation. This suggests that subjects older than 50 have a greater probability of failing to use menu style devices without help. It has to be realized that this study describes an optimistic scenario, as all participants were very motivated at the start of the experiment and higher educated than the average population. Hence, in real life the percentage of drop outs would be even more dramatic.

Slangen-de Kort [1999] categorized the following determinants as important factors in giving up before reaching the goal or persisting until the end: (1) motivation issues that link with needs and goals of the person, (2) personal dispositions concerning the feeling of being able to disengage from options that cannot be achieved anymore and to flexibly readjust one's goals, and (3) perceived competence and control. They refer respectively to the person's perception about his or her personal qualities, and the extent to which he or she thinks to be in charge to influence or produce desired events. The lower the needs, flexibility, and (perceived) competence and control, the sooner the person will leave active adaptation and divert on to the passive strategy.

Within this scheme motivation and flexibility were probably not an issue here: first, people participated voluntarily, which means that they at least started with motivation and they could not choose for alternative paths to achieve the goals, because asking for help and looking in a manual were not permitted. Therefore, only perceived competence seems to be relevant in this study. A closer look at the visuo-spatial reasoning and working memory scores of the persisting subjects versus the drop-outs of the two oldest two age groups did not indicate a significant relation. Obviously, self-perceived competence has more aspects than the two visuo-spatial abilities that we tested.

Table 5.6: Average number of redundant manipulations, navigations and mode errors per step depicted for each age group and subtask of task 1 (first trial only).

age	action	st1	st2	st3	st4	st5	st6
programming							
20-25 y.	man	(10) 6.1	(10)15.8	(10) 0.0	(8)20.3		
	nav	0.0	7.2	0.4	53.5		
	me	0.4	5.4	0.2	9.9		
45-50 y.	man	(10) 0.2	(10)36.4	(9) 6.6	(8) 4.9		
	nav	0.0	18.5	16.4	41.8		
	me	0.2	8.4	3.6	9.0		
55-60 y.	man	(10) 9.5	(9)13.7	(8)10.4	(6) 4.5		
	nav	0.0	20.6	20.4	67.5		
	me	1.2	12.0	6.2	23.7		
70-75 y.	man	(7) 3.0	(7)80.9	(7) 2.7	(5) 1.4		
	nav	0.0	44.1	3.1	14.8		
	me	0.9	21.0	1.1	3.6		
75-80 y.	man	(8)20.0	(7)15.0	(7) 9.1	(6) 1.3		
	nav	0.0	17.7	10.3	14.8		
	me	7.0	11.7	8.3	2.2		
switching							
20-25 y.	man	(10) 1.7	(10)15.2	(10) 3.0	(8) 0.0	(9) 3.1	
	nav	1.6	9.4	11.9	0.0	12.4	
	me	1.1	2.4	2.1	0.0	5.7	
45-50 y.	man	(10) 5.5	(10)22.40	(10) 0.0	(10) 0.0	(10) 6.7	
	nav	3.4	14.5	1.1	0.8	56.3	
	me	0.0	2.8	0.1	0.0	7.9	
55-60 y.	man	(10) 5.8	(10)20.4	(10)14.7	(9) 1.7	(9) 1.6	
	nav	3.0	23.3	17.0	2.8	15.6	
	me	0.0	6.6	7.0	1.1	2.8	
70-75 y.	man	(7) 4.6	(7)32.3	(7) 2.0	(5) 0.2	(3) 0.7	
	nav	8.4	29.3	9.2	0.2	14.0	
	me	1.3	8.6	2.0	0.0	8.0	
75-80 y.	man	(7) 4.0	(7)11.4	(6) 0.0	(6) 0.0	(5) 0.6	
	nav	2.7	20.0	7.5	3.2	10.4	
	me	0.4	11.7	1.0	0.5	1.4	
naming							
20-25	man	(10) 0.0	(10) 0.0	(10) 0.0	(10) 0.0	(10) 0.0	(10) 0.0
	nav	0.4	0.6	0.0	0.0	0.4	9.4
	me	0.0	0.0	0.0	0.0	0.0	1.8
45-50	man	(10) 0.2	(10) 0.0	(10) 0.0	(10) 0.0	(10) 0.0	(10) 0.0
	nav	2.4	1.4	2.4	6.4	4.1	27.9
	me	1.2	0.0	0.5	1.0	0.0	1.6
55-60	man	(10) 0.2	(10) 0.5	(10) 0.0	(10) 0.0	(10) 0.0	(10) 0.9
	nav	2.3	3.2	0.6	9.0	2.3	68.3
	me	0.3	0.3	0.0	0.7	0.5	6.4
70-75	man	(7) 2.0	(7) 0.2	(7) 0.0	(7) 0.9	(7) 0.0	(5) 0.0
	nav	11.2	2.6	0.6	6.7	0.0	28.8
	me	3.9	0.0	0.0	0.0	0.0	2.2
75-80	man	(7) 0.0	(6) 0.0	(6) 0.0	(6) 0.0	(6) 0.0	(4) 0.0
	nav	0.7	4.2	2.0	10.0	11.8	27.3
	me	0.0	1.2	1.3	1.3	2.7	1.5

nav=navigation, man=manipulation, me=mode errors, st1, 2, 3, 4, 5, 6 = step 1, 2, 3, 4, 5, 6. See table 5.5 for the explanation of procedures per step, numbers between brackets represent the number of subjects that participated in the particular subtask.

Table 5.7: Average number of redundant manipulations, redundant navigations and mode errors per step depicted for each age group and subtask of task 3 (first trial only).

age	action	N	step 1
watching channel			
20-25	man	10	7.4
	nav		12.3
	me		0.6
45-50	man	10	7.9
	nav		15.1
	me		7.0
60-65	man	10	54.4
	nav		63.2
	me		16.0
70-75	man	7	30.4
	nav		77.8
	me		7.1
75-80	man	3	17.3
	nav		71.0
	me		21.3
changing volume			
20-25	man	10	1.3
	nav		0
	me		0
45-50	man	10	11.8
	nav		9
	me		3.9
60-65	man	10	12.5
	nav		9.7
	me		9.3
70-75	man	7	8.6
	nav		0
	me		0
75-80	man	3	2.3
	nav		0
	me		0

nav=navigation, man=manipulation, me=mode errors. See table 5.5 for the explanation of procedures per step.

5.4.2 *Efficiency*

Effective subjects of the electro-mechanical generation were less efficient than the menu generation. Relatively more redundant actions were executed and the action duration was on average longer. Within the electro-mechanical generation the action duration increased with age, but the number of redundant actions was higher for the youngest age groups compared to the oldest two groups (see figure 5.7). Hence, the user strategy changed with age from a speedy trial-and-error strategy for the younger groups to a reflective strategy for the older ones. This result resembles earlier findings in chapter 4. Also, all age groups succeeded to learn and were rather efficient after practice, irrespective of their initial strategy, although older age groups remained slower. It may be of interest that the two younger age groups were both still employed, whereas the two older ones had already retired.

5.4.3 *Bottleneck handling*

At the start of the experiment the two youngest age groups of the electro-mechanical generation showed a similar learning strategy as the menu generation: more than 50 percent of their redundant actions consisted of navigations. The older age groups engaged in relatively more redundant manipulations to tackle bottlenecks. Hence, again the employed subjects differs from the retired group. It may be the case that the working environment educated the employees to learn typical menu-style tricks. This idea seems to be supported by Pieters [2000], who indicates that the user's strategy is partially influenced by the environment at work in which he/she is able to practice the skills. Consequently, the employed subjects in this study have been able to employ a speedy trial-and-error strategy, which has been found to be a helpful learning strategy. On the contrary, the initial reflective strategy of the two older age groups consisted of mainly redundant manipulations, which is a less useful learning strategy.

With experience the electro-mechanical generation was able to adapt to a more appropriate learning strategy. The older age groups learned to tackle bottlenecks by almost 90 percent redundant navigations. This extends the results of Westerink, Majoor and Docampo Rama [2000] towards older age groups.

Basic functions can become bottlenecks themselves when activating them after the use of menu functions. Within the electro-mechanical generation a stronger tendency could be seen with increasing age to use a learning strategy based upon redundant navigations, which is a typical strategy to learn menu-style functions. Hence, the persistence of sticking to an inappropriate procedure is initially stronger with increasing age.

5.5 Conclusions

Within the electro-mechanical generation an age-related change in user strategy has been found. The ineffectiveness in terms of the number of drop-outs and failing subjects increased with age. Those subjects who were effective and used an active adaptation strategy were initially inefficient compared to the menu generation. The youngest age groups of the electro-mechanical generation used a speedy trial-and-error strategy in which they tackled menu-related bottlenecks mainly by navigation. On the contrary, the older ones tended to use a reflective strategy in which they initially handled bottlenecks mainly by manipulating. With practice these older age groups were able to switch from redundant manipulations, which is a less helpful learning strategy, towards redundant navigations while dealing with menu-related bottlenecks. With increasing practice, all age groups within the electro-mechanical generation were able to be both more effective and efficient. Temporarily, the electro-mechanical generation persisted in using procedures just learned, even when they were not appropriate anymore.

5.5.1 *Implications for design*

The difficulties of the electro-mechanical generation while controlling menu-style functions suggest that electro-mechanically related (non-layered) procedures, typical basic functions, would be more appropriate for them to use. Furthermore, it has been observed that they had trouble applying different procedures (specialized versus basic ones) on one system. This emphasizes the importance of consistent interfaces that are based upon handling according to one type of procedure. However, when this is not feasible, on-line help could guide users to reach the goals desired. Furthermore, buttons for navigational procedures and manipulations should be physically separated to diminish confusions about the function of each button. Finally, training the (older age groups of the) electro-mechanical generation to learn to use menu-driven interfaces by navigating could help them overcome initial difficulties.

5.5.2 *Automation tool*

Although it has not been the emphasis of this chapter, the importance of the automation tool used in this research is worthwhile to mention. This tool has proven to be valuable, as we have been able to analyse the behaviour patterns in detail. Inevitably, it has produced much more information than we presented in this chapter. It has definitely added value in future applications, for example in applying agent technology to guide the user online. Agents that would be able to register and analyse behaviour patterns like the supporting automation tool in this study could support the user more accurately and customize their help towards his/her need.

6

Conclusions

Technological innovations of the past century have influenced beyond recognition the way our work and private lives are organised. Technology has become a vital element in our daily activities. However, often devices, either in the form of a computer, telephone, video recorder, television, or fax, turned out to be difficult to use. As the technological innovations continue to grow rapidly in the 21st century, it is important that everyone, including older adults, are informed about the new applications of modern technology and how it can be used.

Development of user-friendly interfaces will help users get familiar with the latest technology available in the market. Ultimately the industry will benefit from this as well, as it will boost purchase and usage behaviour.

Development of user-friendly systems requires knowledge about factors that influence the user's learning behaviour while handling new systems. This thesis is concerned with investigating of specific human and user interface characteristics that play a role while handling current electronic consumer products.

6.1 Recapitulation

This thesis focused on the influence of three main factors while handling devices. These factors were concerned with the complexity of the user interface, the age of subjects and the technology generation to which they belong.

In contrast to devices like the bicycle and the stove that have become increasingly easier to use instantly, most interactive electronic consumer products currently available on the market do not have the reputation to be easy to handle. This thesis has focused on the number of interface layers up to four to analyse the complexity of user interfaces as a likely factor for difficulty of use.

Earlier user-system interaction studies showed that difficulties in handling a system increase with age. In general, older users commit more errors and need more time to reach goals (e.g. Czaja [1996]). Czaja and Sharit [1998] found that visuo-spatial working memory plays a role in error performance. A study of Salthouse, Hambrick, Lukas and Dell [1996] indicated that the user's motoric processing speed explains 70 percent of the variance in task duration. Freudenthal, T.D. [1998] found the user's latency was negatively correlated with visuo-spatial working memory, reasoning ability, and processing speed. In this thesis age has been taken into account both separately and in a possible relation to changes in visuo-cognitive abilities.

A third factor that seems to play a role is technology experience. As Freudenthal, A. [1998] found, little children fail to program a television if they have only experienced computer games. Consequently, it may be assumed that older users fail to use present-day technology due to the mismatch between the technology experience they acquired through life and the knowledge they need to use a new device. Sociological studies [Glenn, 1974; Becker, 1992] indicated that the formative period of adolescence and young adulthood is critical for learning norms, values, and skills and these influence behaviour later in life. Subsequent birth cohorts that experienced similar macro-social events during this period may display similar behaviour nowadays. Such cohorts of people can be identified as a generation. In this way Sackmann and Weymann [1994] identified technology generations based on common technological events in the formative period. In this thesis technology generations have been taken into account based on the type of user interface experience they acquired during the formative period.

This final chapter will discuss the main conclusions across chapters concerning the number of interface layers of current user interfaces, age, and generation differences of subjects in handling simulations of present-day consumer electronics products. The chapter will be concluded with some implications for design, instruction, and further research.

6.2 Main results

6.2.1 *User interface complexity*

Consumers and researchers have stressed the point that user interfaces of the early days were easier to use than those currently available on the market, probably due to increase in complexity. Therefore, an overview is needed of user interface changes over the years. At the start of the research there was no study available that had analysed the historical evolution of user interfaces of consumer electronics products, to the best of our knowledge. Therefore, chapter 2 did precisely this, intending to fill this gap. The study showed that user interfaces went through one major change around 1980. User interfaces before 1980 were of the electro-mechanical style and contained only few functionalities for basic goals, robust mechanical buttons that are manually controlled to indicate their state directly, and feature a straightforward interaction based on a one-to-one relation between the button and its function. Many of these characteristics disappeared on software style interfaces of the 1980s (display style) and the 1990s (menu style). Touch buttons without feedback about the state of the device, a central display, software-based objects and many new functions made the interaction symbolic and rather indirect. In these user interfaces a sequence of actions had to be executed to reach the goal, buttons had different functions depending on their mode, and the structure became layered so that part of the information is temporarily hidden. The menu style at least visualizes the layered structure via a menu on a display. At the same time, new types of buttons and procedures have been added to navigate through the menu and activate its functionality. So, the complexity of the user interface increased.

Chapter 3 investigated the impact of interface layers as a major user interface aspect on user performance. The performance of older and younger adults was analysed while handling two user interfaces that did neither differ in functionality nor in input and output devices, but only in the structure (single-layered or two-layered). All subjects turned out more accurate and faster using the single-layered interface than the two-layered user interface. Chapter 3, 4 and 5 showed that the difficulties using a software style increased with its number of layers: significantly more mode errors were made and more time was needed to fulfil the tasks. This result is in agreement with the available data from the literature (e.g. Landauer and Nachbar [1985]).

Chapter 4 searched for some indications of the cognitive skills that are needed to use a software style device. Two skills were studied: reasoning ability for visual objects to identify the logic behind software style procedures, and visuo-spatial working memory aimed at memorizing visuo-spatial positioning of objects or information on a display. Both skills play a significant role. These results corresponded with findings of Czaja and Sharit [1998] and Freudenthal, T.D. [1998]. However, also other factors are expected to play a role that were not captured in this thesis. One of these is advanced hand-eye coordination to perform parallel activities outside the user's visual field. Results at least suggested that this factor impacts the load on visuo-spatial working memory.

Chapter 5 focused on user strategies while dealing with a realistic menu-driven television simulation of four layers. A successful learning strategy to overcome bottlenecks turned out to be the execution of mostly redundant navigations. This was also found by Westerink, Majoor and Docampo Rama [2000]. Such a strategy gives the user the opportunity to get an overview of the structure of the system without manipulating settings that were not meant to be changed. While dealing with menu procedures subjects found it particularly difficult to select the correct option and to reach the final procedure to manipulate the settings. With experience, these bottlenecks disappeared for those who persisted during the first tryout.

6.2.2 *Age and technology generation*

Chapter 2 investigated which technology generations could be identified based on user interface experience in the formative period. A historical analysis, literature research, and an empirical study showed that two main generations can be distinguished: the electro-mechanical generation (born before 1960) and the software generation (born in 1960 or later). This latter generation could be split further into the display generation (born between 1960 and 1970) and the menu generation (born in 1970 or later). These boundaries take into account a diffusion period in which a specific user interface had been introduced on the market, but was not yet adopted and used by the majority of the population.

With this knowledge in mind, a study was set up to analyse age and technology generation effects in handling new systems. Chapter 3 and 4 focused on this research question. As age and generation effects both depend on the year of birth, a sociological method [Glenn, 1977] was chosen to be able to distinguish these effects empirically. An age effect can be identified by a continuous change in performance with age. A generation effect shows a discontinuous change with age. A technology generation effect for user interfaces can be identified if the discontinuity is found around the birth cohort of 1960.

Outcomes of the analyses on 30 to 80 years old indicated that both influences could be identified. An age effect was found for task duration or speed, whereas a technology generation effect was found for error performance. When analysing the data of Czaja [1996] similar effects could be identified. The two effects were found irrespective of the complexity of the device and the user's educational background. Hence, for the first time we have been able to separate age and technology generation effects in handling current user interfaces. The generation effect appeared most strongly during the first confrontation with new user interfaces. With increased experience all age groups were able to handle the user interface more efficiently and the generation effect started to disappear, whereas the age effect continued to be visible, although all older subjects continued to learn.

Chapter 4 focused on two cognitive abilities in relation to the age and technology generation effects found on each performance measure. Outcomes showed that the age effect in task duration relates to two visuo-cognitive skills: reasoning ability and visuo-spatial working memory. This result is in line with findings of Freudenthal, T.D. [1998]. The technology generation effect found in error performance related to the user's visuo-spatial working memory. This result resembled the findings of Czaja and Sharit [1998].

But how could this cognitive factor indicate a generation effect as it is known to decline more or less linearly with age? Detailed analysis of the trends of visuo-spatial working memory in other studies indicated that until the middle of the eighties, visuo-spatial working memory was indeed found to decline linearly with age [Salthouse, 1991; Schaie, 1983]. But recent results concerning this ability suggest a generational shift (see Czaja and Sharit [1998] and chapter 4 of this thesis) that may be based on generation-specific experience with current types of user interfaces (the software style). This was not only found for tests as used in this thesis that resembled working memory problems for software style devices. It was also found for neuro-psychological paper-and-pencil tests like those used by Czaja and Sharit [1998].

But is it possible that generation-specific user interface experience has such a large impact on the user's cognitive skills? Developmental studies showed that the period of adolescence and young adulthood are critical years where experience and cognitive development feed each other constantly for further maturation [Sroufe and Cooper, 1988]. Therefore, an increasing supply of software style devices in the formative period of the software generation may have triggered the increased development of the cognitive skills that are needed to handle such devices. For the software generation the evolution of the user interface may have led to a well trained visuo-spatial working memory. This observation has great implications for the usefulness of traditional visuo-cognitive tests for assessing younger generations. It asks for a serious update of neuro-psychological tests that relate to accumulated experience of the younger generations.

In the studies described in this thesis, all age groups proved to be able to learn to deal with software style devices that had implemented up to three layers (chapter 3 and 4). However, while dealing with a four-layered device that approximated the level of complexity of realistic consumer electronics products also the effectiveness decreased with age, i.e more people of the older age groups dropped out. In real life this is expected to be worse.

Chapters 3, 4 and 5 indicated that the software generation performs both fast and accurately compared to the electro-mechanical generation. Such a strategy resembles that of an expert [Charness, 1981; Singley and Anderson, 1985]. This finding confirms that the software generation has an a priori advantage over the electro-mechanical generation when it concerns using a new software style device.

Only the electro-mechanical generation shows a speed-accuracy tradeoff, that consists of a speedy and inaccurate way of playing around (speedy trial-and-error strategy) by the younger age groups and a cautious and more accurate way of handling (reflective strategy) by the older ones. The choice of the older members of the electro-mechanical generation to slow down to reach the goal implies a simultaneously higher load on visuo-spatial working memory, which is known to decline with age. Chapter 5 showed that such tradeoff also applies to performing tasks of realistically complex systems. At the start this reflective strategy induced many redundant actions that were based upon direct manipulations of the system. Such an approach was functional for electro-mechanical devices, but is not appropriate anymore for devices of today. With experience, also the persisting older age groups of the electro-mechanical generation were able to change their learning strategy towards the execution of redundant navigations as a useful strategy for handling realistically complex menu-style devices.

6.3 Practical implications

The knowledge acquired in this thesis research as part of the Philips Technology Generation project suggests that designers of user interfaces need to be aware of the generation-specific ability gap between potential users born before 1960 and those born later [Docampo Rama, 2000]. Ease-of-use of software style devices could be improved by re-introduction of electro-mechanical style concepts such as the one-to-one relation between function and button, understandable direct feedback about the state of the device and increase of affordance (i.e. the user's perception about the applicability of the device) [Norman, 1988], for a straightforward interaction. This will reduce the load on a number of cognitive skills that are known to decline with age.

Reduction of the number of layers could be one such improvement. The fact that even some of the higher educated older adults gave up after being confronted with a four-layered interface strongly indicates the need to reduce the number of layers. But also aspects that provoke inconsistent spatial placing of functions should be dealt with, for example eliminating laborious procedures, multi-functional buttons, and mode changes. Diverse input and output devices such as the wireless remote control and mouse-pointing device hamper a smooth hand-eye coordination, putting even more load on the user's cognitive abilities.

Other ways to reduce the complexity of software-style user interfaces could consist of providing more embedded guidance to the user and having complicated procedures taken over by the system itself instead of bothering the user with it. Adaptive technology [Masthoff, 1997] can be implemented to assist the user on-line. Intelligent agents identify the user's need by focussing on the user's navigation behaviour. Also the user's average duration per action is a simple and accurate predictor of user difficulties, which holds across generations and is particularly successful within the electro-mechanical generation. Situated learning guided by an agent could help them overcome bottlenecks. Online agents that are able to register and analyse behaviour patterns like the automation tool in this study could support users more accurately and tailor its help towards their needs.

Training the electro-mechanical generation could also be useful. It is then recommended to teach older users to explore rather than being cautious. Learning to first explore the user interface by navigational actions to get an overview of the structure of the device is a must here. Furthermore, speeding up their action rate may reduce the load on their visuo-spatial working memory. This will facilitate them to develop a mental model that matches with the way a software style device can be made to work. Generation-specific (on-line) training could focus on the logic behind software style concepts and procedures to activate specialized functionality (e.g. programming functions, services). This may reduce the reasoning problems of the electro-mechanical generation while using software style devices. Online help or instructions could focus on explanations on how to find the correct option and how settings can be manipulated (final step of each goal). This type of instruction would be effective for all generations.

Optimal use of such guidelines in product development can be achieved by involving usability specialists and consumer researchers from the start of the Product Creation Process [Docampo Rama, 1998]. Transfer of relevant information to designers and engineers can then best be judged and adapted to the possibilities and restrictions of the implementation. The true key of success of such multi-disciplinary team is based on inter-mutative learning: putting effort in understanding and learning from each others work and approach, which implies an intensive cooperation.

6.4 Implications for further research

Challenges in this new millennium for further research on ageing and user-system interaction can be found in (1) further analysis of technology generation effects for a deeper understanding of its background in earlier experience and of its implications, (2) studies on the implications of new types of user interfaces that enter the market as to implications on present and future technology generations, (3) integration of age and technology generations in the development of new systems, user interfaces, and on-line and off-line training programs, and (4) revisiting of neuro-psychological tests to take into account evolving experiences in the daily handling of technology. This will enable the assessment of (future) technology generations on their evolving visuo-cognitive and other capabilities.

References

ANDERSON, J.R. [1983]. *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.

ANDERSON, J.R. [1990]. *Cognitive Psychology and its Implications*. New York: W.H. Freeman and Company.

BADDELEY, A.D. [1986]. *Working Memory*. Oxford: Clarendon Press.

BADDELEY, A.D., AND HITCH, G.J. [1974]. Working memory. *The Psychology of Learning and Motivation*, **8**, 47–90.

BALTES, P.B., AND BALTES, M.M. [1990]. Psychological perspectives on succesful aging: The model of selective optimization with compensation. In P.B. Baltes and M.M. Baltes (Eds.), *Successful Aging: Perspectives from the behavioral sciences*. New York: Cambridge University Press.

BAUDET, H. [1986]. *Een vertrouwde Wereld: 100 jaar innovatie in Nederland*. Amsterdam: Bakker.

BAXTER, I., AND OATLEY, K. [1991]. Measuring the learnability of spreadsheets in inexperienced users and those with previous spreadsheet experience. *Behaviour and Information Technology*, **10**, 475–490.

BECKER, H.A. [1992]. *Generaties en hun Kansen*. Amsterdam: Meulenhoff.

BECKER, H.A. [2000]. Discontinuous change and generational contracts. In S. Arber and C. Attias-Donfut (Eds.), *The myth of generational conflict*, 114–132. England: Routledge & Keegan Paul.

BECKER, H.A., AND GOOR, A. VAN DE [1997]. *Key components of the model explaining the emergence and development of general generations*. Report 8. Department of Sociology, Utrecht University. Utrecht, The Netherlands.

BIRREN, J.E., AND SCHAIE, K.W. [1990]. *The Handbook of Psychology of Aging. Third edition*. San Diego, California: Academic Pres, Inc.

BOUMA, J.M, LINDEBOOM, J., AND HOUTE, L. VAN [1990]. *Klinisch Neuropsychologische Assessment. Een handleiding voor de praktijk*. Nijmegen: KUN.

BRANDTSTÄDTER, J., AND RENNER, G. [1990]. Tenacious goal pursuit and flexible goal adjustment: Explanation and age-related analysis of assimilation and accommodation strategies of coping. *Psychology and Aging*, **5**, 58–67.

BRINKMAN, W.P. [2000]. *Human Behaviour and the Comprehension of Interaction Components*. Report. IPO: Center for User-System Interaction. Eindhoven, The Netherlands.

BURKE, D.M. [1985]. Raven's progressive matrices (1938): More on norms, reliability, and validity. *Journal of Clinical Psychology*, **41**, 231–235.

CERELLA, J., POON, L.W., AND WILLIAMS, D.M. [1980]. Age and the complexity hypothesis. In L.W. Poon (Ed.), *Aging in the 80s. Psychological issues*, 332–340. Washington, DC: American Psychological Association.

CHARNESS, N. [1981]. Search in chess: Age and skill differences. *Journal of Experimental Psychology: Human perception and performance*, **7**, 467–476.

COLBERT, M. [1997]. Style guides and their application: the case of Microsoft "Windows" and a remote tutoring environment. *Behaviour and Information Technology*, **16**, 25–42.

CONSUMENTENBOND [1953 to 1997]. Consumentengids. *Monthly Journal of the Consumentenbond*, Heerlen, The Netherlands.

CORKIN, S. [1983]. Some relationships between global amnesias and the memory impairments in alzheimer's disease. In S. Corkin, K.L. Davis, J.H. Growdon, E. Isdin, and R.J. Wiertsma (Eds.), *Alzheimer's Disease: A report of progressive aging, volume 19*. New York: Raven Press.

CORSI, P.M. [1972]. *Human Memory and the medial temporal Region of the Brain*. Ph.D. thesis. McGill University.

CRAIK, F.I.M., AND BOSMAN, E.A. [1992]. Age-related changes in memory and learning. In H. Bouma and J.A.M. Graafmans (Eds.), *Gerontechnology*. Amsterdam: IOS Press.

CRAIK, F.I.M., MORRIS, R.G., AND GICK, M.L. [1990]. Adult age differences in working memory. In G. Vallar and T. Shallice (Eds.), *Neuropsychological Impairments of Short-Term Memory*, 247–267. Cambridge: Cambridge University Press.

CUTLER, S.J., AND KAUFMAN, R.L. [1975]. Cohort changes in political attitudes; tolerance of ideological nonconformity. *Public Opinion Quarterly*, **2**, 68–81.

CZAJA, S. [1996]. Aging and the acquisition of computer skills. In W.A. Rogers, A.D. Fisk, and N. Walker (Eds.), *Aging and Skilled Performance: Advances in theory and applications*, 201–220. Mahwah, NJ: Lawrence Erlbaum Associates.

CZAJA, S.J., AND SHARIT, J. [1993]. Age differences in the performance of computer-based work. *Psychology and Aging*, **8**, 59–67.

CZAJA, S.J., AND SHARIT, J. [1998]. Ability-performance relationships as a function of age and task experience for a data entry task. *Journal of Experimental Psychology: Applied*, **4**, 332–351.

DE RENZI, E., FAGLIONI, P., AND PREVIDI, P. [1977]. Spatial memory and hemispheric locus of lesion. *Cortex*, **13**, 424–433.

DETWEILER, M.C., HESS, S.M., AND ELLIS, R.D. [1996]. The effects of display layout on keeping track of visuo-spatial information. In W.A. Rogers, A.D. Fisk, and N. Walker (Eds.), *Aging and Skilled Performance: Advances in theory and applications*, 157–184. Mahwah, NJ: Lawrence Erlbaum Associates.

DOCAMPO RAMA, M., AND KAADEN, F. VAN DER [2001]. Technology generations: Characterization on the basis of user interface changes. *Ageing and Technology: Starting into the third Millenium. Third international conference of gerontechnology*, Accepted. München: Technische Universität München.

DOCAMPO RAMA, M. [1997]. Age-related learning effects in working with layered interfaces. *IPO Annual Progress Report*, **32**, 19–26.

DOCAMPO RAMA, M. [1998]. Learnability of consumer electronics. In M. Sikorski and M. Rauterberg (Eds.), *Transferring Usability Engineering to Industry*, 62–64. Gdańks: Technical University of Gdańsk.

DOCAMPO RAMA, M. [2000]. Waarom moderne apparaten lastiger te bedienen zijn voor de oudere generatie gebruikers. *Géron*, **6**, 73–79.

DOCAMPO RAMA, M., AND KAADEN, F. VAN DER [1998]. *The User-Interface of Consumer Electronics: Historical overview of interaction styles in the Netherlands, 1963-1996*. Report no. 1194. IPO: Center for Research on User-System Interaction. Eindhoven, The Netherlands.

FITTS, P.M. [1964]. Perceptual-motor skill learning. In A.W. Melton (Ed.), *Categories of Human Learning*. New York: Academic Press.

FREUDENTHAL, A. [1998]. Transgenerational design of “smart products”, a checklist of guidelines. In J. Graafmans, V. Taipale, and N. Charness (Eds.), *Gerontechnology A Sustainable Investment in the Future. II: Case studies*, 406–410. Amsterdam: IOS Press.

FREUDENTHAL, T.D. [1998]. *Learning to Use Interactive Devices; Age differences in the reasoning process*. Ph.D. thesis. Eindhoven University of Technology.

GLENN, N.D. [1974]. Aging and conservatism. *Annuals of the American Academy of Political and Social Science*, **11**, 176–186.

GLENN, N.D. [1977]. Cohort analyses. In E.M. Uslaner (Ed.), *Series: Quantitative Applications in the Social Sciences*. Beverly Hills, California: SAGE Publications inc.

GOMEZ, L., EGAN, D., WHEELER, E., SHARMA, D.K., AND GRUCHACZ, A. [1983]. How interface design determines who has difficulty learning to use a text editor. *Proceedings of CHI'83 Human Factors in Computing Systems*, 176–181. New York: Association of Computing Machinery.

GOOR, A.G. VAN DE [1998]. Generaties in techniekbezit: eerste resultaten van het philips techniekgeneraties-project. In P. Hermkens, R. van Rijsselt, and K. Sanders (Eds.), *Differentiatie en samenleving: Opstellen voor Henk Becker*, 101–130. Amsterdam: Thela Thesis.

GOOR, A.G. VAN DE, AND BECKER, H. [2001]. *Technology Generations in the Netherlands, a sociological Analysis*. Maastricht: Shaker.

HAAKMA, R. [1998]. *Layered Feedback in User-System Interaction*. Ph.D. thesis. Eindhoven University of Technology.

HALE, S., BRONIK, M.D., AND FRY, A.F. [1997]. Verbal and spatial working memory in school-age children: Developmental differences in susceptibility to interference. *Developmental Psychology*, **33**, 364–371.

HANSMAN, H. [1996]. *Life Course Experience and Music Consumption. An analysis of the adoption of the compact disc player and purchase of compact discs in the Netherlands*. Ph.D. thesis. Amsterdam.

HERON, A., AND CHOWN, S.M. [1967]. *Age and Function*. Boston: Little-Brown & Co.

HOGESTEEGER, G. [1989]. *Van lopende Bode tot Telematica*. Groningen: PTT.

HOWARD, J.H., AND HOWARD, D.V. [1997]. Learning and memory. In A.D. Fisk and W.A. Rogers (Eds.), *Handbook of Human Factors and the older Adult*, 7–26. San Diego, California: Academic Press, Inc.

HULTSCH, D.F., AND DIXON, R.A. [1990]. Learning and memory and aging. In J.E. Birren and K.W. Schaie (Eds.), *Handbook of the Psychology of Aging (3rd ed.)*. New York: Van Nostrand Reinhold.

HUTCHINS, E.L., HOLLAN, J.D., AND NORMAN, D.A. [1985]. Direct manipulation interfaces. *Human-Computer Interaction*, **1**, 311–338.

ISO [1993]. *Guidance on specifying and measuring usability*. ISO CD 9241 part 11: draft.

JANSEN, W., AND WITTENBOER, G.L.H. VAN DEN [1992]. *Sociale Netwerken en hun Invloed*. Amsterdam: Boom.

JOHNSON, J. [1990]. Modes in non-computer devices. *International Journal of Man-Machine Studies*, **32**, 423–438.

KEATING, D.P. [1980]. Thinking processes in adolescence. In J. Adelson (Ed.), *Handbook of Adolescent Psychology*. New York: Wiley.

KELLEY, C.L. [1996]. *Age, Experience, and Transfer across Word Processors*. Ph.D. thesis. University of Waterloo.

LANDAUER, T. K., AND NACHBAR, D.W. [1985]. Selection from alphabetic and numeric menu trees using a touch screen: Breadth, depth and width. In L. Borman and B. Curtis (Eds.), *Human Factors in Computing Systems II*, 73–78. Amsterdam: Elsevier Science Publishers.

LAWTON, M.P. [1975]. Environmental proactivity in older people. In V.L. Bengtson and K.W. Schaie (Eds.), *The Course of later Life: Research and reflections*, 15–23. New York: Springer.

LAZARUS, R.S., AND FOLKMAN, S. [1984]. Coping and adaptation. In W.D. Gentry (Ed.), *Handbook of Behavioral Medicine*, 282–325. New York: Guilford.

LOGIE, R.H. [1995]. *Visuo-spatial Working Memory*. Hove: Lawrence Erlbaum Associates.

MASTHOFF, J.F.M. [1997]. *An Agent-Based Interactive Instruction System*. Ph.D. thesis. Eindhoven University of Technology.

MELENHORST, A.S. [2001]. Older people planning future holidays: The role of time. *Ageing and Technology: Starting into the third Millenium. Third international conference of gerontechnology*, Accepted. München: Technische Universität München.

MILNER, B. [1971]. Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin*, **27**, 272–277.

NEIMARK, E.D. [1982]. Adolescent thought: Transition to formal operations. In B.B. Wolman and G. Strickler (Eds.), *Handbook of Developmental Pyschology*. Englewood, NJ: Prentice-Hall.

NORMAN, D.A. [1988]. *The Psychology of Everyday Things*. New York: Basic Books.

PIAGET, J., AND INHELDER, B. [1973]. *Memory and Intelligence*. New York: Basic Books.

PIETERS, J.M. [2000]. *Cognitieve Veroudering: Een kwestie van informatieverwerking, kennisontwikkeling en aanpassing*. Report. University of Twente. Twente, The Netherlands.

RAUTERBERG, M. [1992]. An empirical comparison of menu-selection (cui) and desktop (gui) computer programs carried out by beginners and experts. *Behaviour and Information Technology*, **4**, 227–236.

RAVEN, J.C. [1960]. *Guide to the Standard Progressive Matrices*. London: H.K. Lewis.

REASON, J. [1990]. *Human Error*. New York: Cambridge University Press.

REITAN, R.M. [1980]. Validity of the trail making test as an indication of organic brain damage. *Perceptual and Motor Skills*, **6**, 271–282.

ROGERS, A., AND FISK, D. [2000]. Human factors, applied cognition, and aging. In F.I.M. Craik and T.A. Salthouse (Eds.), *Handbook of Aging and Cognition. Second edition*, 559–592. Mahwah, NJ: Lawrence Erlbaum Associates.

ROGERS, M. [1983]. *Diffusion of Innovations (3rd ed.)*. New York: Free Press.

RUBIN, D.C., RAHHAL, T.A., AND POON, L.W. [1998]. Things learned in early adulthood are remembered best. *Memory & Cognition*, **26**, 3–19.

SACKMANN, R., AND WEYMANN, A. [1994]. *Die Technisierung des Alltags: Generationen und technische Innovationen*. Bremen: BMFT.

SALTHOUSE, T.A. [1991]. *Theoretical Perspectives on Cognitive Aging*. Hillsdale, NJ: Lawrence Erlbaum Associates.

SALTHOUSE, T.A., HAMBRICK, D.Z., LUKAS, K.E., AND DELL, T.C. [1996]. Determinants of adult age differences on synthetic work performance. *Journal of Experimental Psychology: Applied*, **2**, 305–329.

SALTHOUSE, T.A., AND SKOVRONEK, E. [1992]. Within-context assessment of age differences in working memory. *Journal of Gerontology: Psychological Science*, **47**, 110–120.

SCAPIN, D.L., AND PIERRET-GOLBREICH, C. [1990]. Towards a method for task description: Mad. In L. Berlinguet and D. Berthelette (Eds.), *In Work with Display Units 89*. North-Holland: Elsevier.

SCHAIE, K.W. [1983]. The seattle longitudinal study: A 21-year exploration of psychometric intelligence in adulthood. In K.W. Schaie (Ed.), *Longitudinal studies of adult psychological development*, 64–135. New York: Guilford Press.

SCHAIE, K.W. [1985]. *Manual for the Schaie-Thurstone Adult Mental Abilities Test (STAMAT)*. Palo Alto, CA: Consulting Psychologists Press.

SCHNEIDERMAN, B. [1983]. *Direct Manipulation: A step beyond programming languages*. Ph.D. thesis. University of Maryland.

SCHOT, J.W., LINTSEN, H.W., RIP, A., AND BRUHÉZE, A.A.A. DE LA [2000]. *Techniek in Nederland in de twintigste Eeuw. I: Techniek in ontwikkeling, waterstaat, kantoor en informatietechnologie*. Stichting Historie der Techniek Walburg Pers.

SCHRAM, M.J. [1998]. *TVSIM: A software-simulated television for testing the learnability of multi-layered user interfaces*. Report. IPO: Center for User-System Interaction. Eindhoven, The Netherlands.

SEARLEMAN, A., AND DOUGLAS, H. [1994]. *Memory from broader Perspective*. New York: McGraw-Hill.

SEBILLOTTE, S. [1995]. Methodology guide to task analysis with the goal of extracting relevant characteristics for human-computer interfaces. *International Journal of Human-Computer Interaction*, **7**, 341–363.

SEBILLOTTE, S., AND SCAPIN, D.L. [1994]. From user's task knowledge to high-level interface specification. *International Journal of Human-Computer Interaction*, **6**, 1–15.

SINGLEY, M.K., AND ANDERSON, J.R. [1985]. The transfer of text-editing skill. *International Journal of Man-Machine Studies*, **22**, 403–423.

SLANGEN-DE KORT, Y.A.W. [1999]. *A Tale of Two Adaptations: Coping processes of older persons in the domain of independent living*. Ph.D. thesis. Eindhoven.

SLATER, P. [1947]. The association between age and score in the progressive matrices test. *Journal of Psychology, Statistical Section*, **1**, 64–69.

SROUFE, L.A., AND COOPER, R.G. [1988]. *Child Development: Its nature and course*. New York: Alfred A. Knopf, Inc.

STEWART, T. [1992]. Physical interfaces or "obviously it's for the elderly, it's grey, boring and dull". In H. Bouma and J.A.M. Graafmans (Eds.), *Gerontechnology*. Amsterdam: IOS Press.

TAYLOR, M.M. [1988a]. Layered protocols for computer-human dialogue. I: principles. *International Journal of Man-Machine Studies*, **28**, 175–218.

TAYLOR, M.M. [1988b]. Layered protocols for computer-human dialogue. II: some practical implications. *International Journal of Man-Machine Studies*, **28**, 219–257.

TAYLOR, M.M. [1989]. Response timing in layered protocols: A cybernetic view of natural dialogue. In M.M. Taylor, F. Néel, and D.G. Bouwhuis (Eds.), *The Structure of Multimodal Dialogue*, 159–172. Amsterdam: North-Holland.

TIFFIN, J. [1968]. *Purdue Pegboard Examiner's Test*. Rosemont, IL: London House.

TOURS, J.D. [1981]. Als het maar zwart is. de ontwikkeling van het gewone telefoontoestel. In J.H. Schuilenga, J.D. Tours, J.G. Visser, and J. Bruggeman (Eds.), *Honderd Jaar Telefoon. Geschiedenis van de openbare telefonie in Nederland 1881-1981*. Den Haag: PTT.

VERNON, P.E. [1947]. The variations of intelligence with occupation, age and locality. *Journal of Psychology Statistical Section*, **1**, 52–63.

VET, J. DE, AND RUYTER, B. DE [1996]. *An Inventory of Interaction Styles*. Report no. 1113. Institute for Perception Research (IPO). Eindhoven, The Netherlands.

WECHSLER, D. [1955]. *Manual for the Wechsler Scale*. New York: The Psychological Corporation.

WECHSLER, D. [1981]. *Manual for the Wechsler Adult Intelligence Scale - Revised*. New York: The Psychological Corporation.

WECHSLER, D. [1987]. *Wechsler Memory Scale*. New York: The Psychological Corporation.

WELFORD, A.T. [1958]. *Ageing and Human Skill*. London: Oxford University Press.

WELFORD, A.T. [1968]. *Fundamentals of Skill*. London: Methuen & CO. LTD.

WELFORD, A.T. [1980]. Memory and age: A perspective view. In L.W. Poon, J.L. Fozard, L.S. Cermak, D. Arenberg, and L.W. Thompson (Eds.), *New Directions in Memory and Aging*, 1–17. Hillsdale, NJ: Lawrence Erlbaum Associates.

WESTERINK, J.H.D.M., MAJOOR, B.G.M.M., AND DOCAMPO RAMA, M. [2000]. Interacting with infotainment applications: navigation patterns and mental models. *Behaviour & Information Technology*, **19**, 97–106.

WIEDENBECK, S., AND DAVIS, S.A. [1997]. The influence of interaction style and experience on user perceptions of software packages. *Journal of Human-Computer Studies*, **46**, 563–588.

WINER, B.J., BROWN, D.R., AND MICHELS, K.M. [1991]. *Statistical Principles in experimental Design (Third edition)*. New York: McGraw-Hill Inc.

WISTER, A.V. [1989]. Environmental adaptation by persons in their later life. *Research on Aging*, **11**, 267–291.

Samenvatting

Ouders reageren wel eens verbaasd als hun kinderen in een mum van tijd de muis van de computer weten te bedienen, computerspelletjes beheersen, of de videorecorder kunnen programmeren. Ze herinneren zich op zulke momenten misschien hoe moeilijk zij het vonden toen ze er zelf mee begonnen. Steeds vaker zijn het de kleinkinderen die opa en oma uitleggen hoe ze de mobiele telefoon moeten gebruiken, de video kunnen programmeren en met de computer kunnen omgaan. En wat te denken van alle opschudding die ontstaat wanneer een kabelmaatschappij het weer eens in zijn hoofd haalt om alle frequenties van de televisiekanalen te veranderen.

Het is altijd zo geweest dat levenservaringen oudere mensen in hun beroep en in sociaal opzicht hebben doen rijpen. Jongeren konden veel van hen leren. Daarom was tot op heden de leraar-leerling relatie tussen ouderen en jongeren eenduidiger: de oudere leert aan de jongere. Echter, naar aanleiding van de snel opeenvolgende technologische ontwikkelingen in de tweede helft van de twintigste eeuw is deze relatie ambigu geworden. In situaties waar recent ontwikkelde technologiekennis gebruikt wordt, zijn het vaak de jongeren die ouderen instrueren.

De technologische ontwikkelingen van de twintigste eeuw zijn ver doorgedrongen in de maatschappij: in het huishouden, op het werk, in financiële zaken, bij medische behandelingen. Het is dan ook van wezenlijk belang dat alle volwassenen, en dus ook ouderen, leren omgaan met moderne apparatuur. Technologie moet immers bijdragen aan iemands wensen van comfort en zelfredzaamheid. Alleen dan kan ieder voor zichzelf uitmaken of een bepaald stukje technologie een handig hulpmiddel is in het realiseren van zijn of haar wensen en behoeften.

Het blijkt dat oudere mensen nogal eens problemen ervaren bij het gebruik van hedendaagse apparatuur. Waarom precies dat zo is, hebben wij nagegaan in een reeks studies binnen het project TechniekGeneraties dat is opgezet en gecoördineerd door Philips Design.

Veel ouderen merken vaak op dat apparaten tegenwoordig veel moeilijker te gebruiken zijn dan vroeger. In dit proefschrift wordt nagegaan in hoeverre dit komt omdat (1) de bediening van apparaten in de loop der tijd complexer is geworden, (2) cognitieve veranderingen door de jaren heen er voor zorgen dat ouderen meer moeite hebben met het leren bedienen van nieuwe apparaten, en (3) ouderen niet opgegroeid zijn met de huidige bedieningsstijl, en daarmee kennis hebben opgebouwd in hun jeugd en adolescentie die niet overeenkomt met de kennis die nodig is om huidige apparatuur te bedienen.

Deze drie hypothesen kunnen getoetst worden door jongeren en ouderen geheel nieuwe apparaten te laten gebruiken, het moeilijkheden te observeren en per apparaat het verloop met de leeftijd te analyseren. Indien aparaten inderdaad complexer zijn geworden zullen zowel jongeren als ouderen minder moeite moeten hebben met apparaten die kenmerken hebben van vroegere apparaten (bv. met enkelgelaagde interactiestructuur), dan met hedendaagse apparaten. Indien cognitieve veranderingen met de leeftijd een rol spelen zullen de moeilijkheden continu met de leeftijd toenemen (leeftijdseffect). Echter als het opgroeien met een bepaalde bedieningsstijl een stempel drukt op het probleemgedrag, dan zal dit gedrag discontinu met de leeftijd oplopen (generatie-effect). De discontinuitiet zou daarbij overeen moeten komen met de verschillende generaties die te onderscheiden zijn op basis van de bedieningsstijl waar ze mee zijn opgegroeid.

Alvorens deze laatste hypothese tesamen met de overigen te kunnen toetsen, is een historische analyse nodig van de ontwikkeling van de bediening van consumenten electronica over de afgelopen eeuw. Hiervoor zijn de televisie, telefoon en videorecorder als voorbeeld genomen. Hoofdstuk 2 beschrijft deze analyse, stelt de bedieningsstijlen vast en berekent de generatieovergangen. Met behulp van een landelijke enquête (1015 mensen) zijn de generatiegrenzen getoetst en bevestigd. Er kunnen twee hoofdgeneraties onderscheiden worden: de electromechanische generatie (geboortecohorten 1959 en eerder) en de softwaregeneratie (geboortecohorten 1960 en later). Deze laatste generatie kan nogeens onderverdeeld worden in de displaygeneratie (geboortecohorten 1960-1969) en de menugeneratie (geboortecohorten 1970 en later).

Hoofdstuk 3 toetst de drie gestelde hypothesen door twee simulaties aan te bieden van een beeldtelefoon die verschillen in gelaagdheid (1 en 2 lagen). Het enkelgelaagd systeem heeft zo een kenmerk van vroegere (electro-mechanische) bedieningspanelen. De simulaties zijn in twee sessies aan vier leeftijdsgroepen aangeboden. De jongste groep is tussen de 25 en 35 jaar, geboren na 1960 en daarmee behorend tot de softwaregeneratie. De overige groepen zijn 40-50, 50-60 en 65-75 jaar en behoren allen tot de electro-mechanische generatie. De bedieningsproblemen werden gemeten aan de hand van de tijdsduur en het aantal fouten.

De resultaten laten zien dat iedereen leert, ook ouderen. Het dubbelgelaagd systeem blijkt voor iedereen, en dus ook voor jongeren, moeilijker te bedienen.

De maat waarmee het probleemgedrag is gemeten bepaalt of er een leeftijdseffect of een generatie-effect te zien is. De tijdsduur loopt continu met de leeftijd op, wat duidt op de invloed van cognitieve veranderingen met de jaren. Het foutenaantal verloopt discontinu met de leeftijd, waarbij de discontinuiteit optreedt tussen de groep van 25-35 jaar en die van 40-50 jaar. Dit duidt op een generatie-effect dat zijn oorsprong kent in de bedieningsstijl waarmee men is opgegroeid. Binnen de electro-mechanische generatie is er geen oplopende foutenaantal geconstateerd, wat een puur generatie-effect suggereert.

Het experiment is herhaald voor een ander en complexer apparaat, een simulatie van een mobiele telefoon met 3 lagen. Ditmaal is het opleidingsniveau van de vrijwilligers hoog (mbo/hbo). Ook nu weer wordt er een generatie-effect gevonden voor het aantal fouten en een leeftijdseffect voor de tijdsduur.

Hoofdstuk 4 bestudeert de gevonden trends in de tijdsduur en foutenaantal in relatie tot enkele vaardigheden van de gebruiker. Hierbij is de simulatie met de mobiele telefoon herhaald bij 88 mensen tussen de 25 en 80 jaar met een hoog opleidingsniveau. Van elke gebruiker zijn het visueel ruimtelijk werkgeheugen, het redeneringsvermogen voor visueel materiaal, en de motorische vaardigheid getest, het werktempo gemeten en de opgedane ervaring met consumenten electronica nagevraagd. Resultaten laten zien dat foutenaantal is gerelateerd aan visueel ruimtelijk werk geheugen in relatie tot geavanceerde oog-hand coördinatie. De software generatie weet hiermee een aanpak van een expert te ontwikkelen. Daarentegen gebruikt de electro-mechanische generatie een aanpak van een leek. Binnen deze generatie worden verschilllende gebruikersstrategieën gehanteerd: de jongeren onder deze oudere generatie neigen snel en inaccuraat te handelen, terwijl de ouderen van deze groep een langzamere maar meer accurate handelswijze verkiezen.

Hoofdstuk 5 gaat dieper in op de strategie van de electro-mechnische generatie tijdens het bedienen van een apparaat. Dit keer gaat het om een simulatie van een televisie waarvan de bediening behoort tot de menustijl. Vijf leeftijdsgroepen (20-25 als referentie groep en maakt deel uit de menu generatie, 45-50, 55-60, 70-75 en 75-80 jaar) werden bestudeerd. De gebruikersstrategie is nagegaan door allereerst te kijken naar de effectiviteit van gebruikers (in hoeverre zijn ze in staat taken te volbrengen). Van de gebruikers die de taken wisten te volbrengen (en dus een actieve strategie toepassen) is bekeken hoe efficient ze waren in termen van de tijd die ze erover deden en hoeveel stappen ze nodig bleken te hebben om de taak te volbrengen. Als laatste is gekeken naar de struikelblokken die ze onderweg ontmoeten.

De resultaten laten zien dat de effectiviteit binnen de electro-mechanische generatie met het de leeftijd daalt, waarbij het met name gaat om de oudste group

(75-80 jaar). Van de effectieve personen uit de electro-mechanische generatie blijken de jongere groepen inderdaad een snelle trial-and-error strategie toe te passen waarin ze met behulp van redundante navigaties op een succesvolle wijze de struikelblokken weten te overwinnen. De effectieve ouderen uit diezelfde generatie prefereren aanvankelijk een reflectieve strategie, waarbij ze hindernissen aanpakken met behulp van redundante manipulaties, wat leidt tot minder snelle afhandeling van struikelblokken. Maar ook zij leren met behulp van redundante navigaties de problemen op te lossen. Typische problemen bij menu's blijken het selecteren van de juiste keuze en de overgang van navigaties naar manipulaties.

Het proefschrift wordt beëindigd met hoofdstuk 6 waarin conclusies worden getrokken over de hoofdstukken heen, en nagedacht tot welke ontwerprichtlijnen de resultaten hebben geleid. Terugkijkend op het onderzoek beschreven in dit proefschrift kan worden geconcludeerd dat bedieningspanelen inderdaad moeilijker te gebruiken zijn naarmate het uit meer lagen bestaat. De moeilijkheden nemen met de leeftijd toe, als het gaat om de tijd die men nodig heeft om het apparaat te bedienen. Daarnaast blijken het aantal fouten en de strategie generatiespecifiek te zijn, en dan metname bij de eerste confrontatie met een nieuw apparaat. Aangezien de meeste beslissingen tijdens zo'n eerste indruk worden genomen is het daarom zo van belang om het generatie-specifieke referentiekader van de consument of gebruiker in acht te nemen bij het aanpassen van het bedieningspaneel. Het trainen-op-maat behoort ook tot de mogelijkheden waarin het hiaat tussen het referentiekader van de gebruiker en de nieuwe procedures gedicht wordt.

Curriculum Vitae

Maria de los Milagros Docampo Rama was born in Tegelen, the Netherlands on 18 February 1972. She attended the Blariacum College at Venlo-Blerick from 1984 to 1990, where she obtained her VWO diploma in 1990. From 1990 to 1996 she studied Psychology at the University of Nijmegen. She obtained her Master's degree in psychology with honor in April 1996, specializing in neuropsychology and 18 additional courses in clinical and cognitive psychology. Her master thesis was entitled 'audio-visual information processing' and won the 1996 thesis award of the NVN (Dutch association for Neuropsychology). From May 1996 to May 2000 she was employed by the former institute of Gerontechnology at IPO Center for User-System Interaction at the Eindhoven University of Technology. During this period she worked for Technology Generations, a project commissioned by Philips Design. The work resulted in this Ph.D. thesis. Since May 2000 she is working at Philips Design as a consumer research consultant within Strategic Design. Apart from the project coordination of Technology Generations, she is working on trend and consumer research and e-business.