Stress in senior computer interaction

DOI:
10.6100/IR633490

Document status and date:
Published: 01/01/2008

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Stress in Senior Computer Interaction

Henk Herman Nap

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

Printing: Universiteitsdrukkerij Technische Universiteit Eindhoven
Cover realisation: Henk Herman Nap
An electronic copy of this thesis is available from the site of the Eindhoven University Library in PDF format (http://www.tue.nl/bib).

CIP-DATA LIBRARY TECHNISCHE UNIVERSITEIT EINDHOVEN

Nap, Henk Herman

Stress in senior computer interaction / by Henk Herman Nap. - Eindhoven : Technische Universiteit Eindhoven, 2008. – Proefschrift. - ISBN 978-90-386-1227-0
NUR 778
Keywords: Stress / Seniors / Elderly / System complexity / Interface design / Performance feedback / Gerontechnology

Stress in Senior Computer Interaction

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
Rector Magnificus, prof.dr.ir. C.J. van Duijn, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen
op dinsdag 18 maart 2008 om 16.00 uur

door

Henk Herman Nap

geboren te Hardenberg

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr. D.G. Bouwhuis
en
prof.dr. M.A. Neerincx

Copromotor:
dr. H.P. de Greef

Acknowledgements

First I thank Paul de Greef for support, guidance and Arabica coffee. He helped me to create order in thoughts and documentation. I am grateful to Don Bouwhuis for the valuable comments on my work, for sharing his ideas about Gerontechnology, and being there when needed. Thanks to Mark Neerincx for carefully reviewing my thesis.

I would also like to thank Jim Juola, Cees Midden, and Bertie Zwetsloot-Schonk for participating in my committee and for their constructive comments on the thesis. In addition, thanks to Berry Eggen for taking part in the promotion committee.

I also thank all the colleagues at HTI for providing a unique and inspiring work atmosphere. Special thanks go to Martin Boschman for technology and lab support. Furthermore, thanks to Vera Diebels and Anita Nellissen for invaluable social, office, and administrative support.

I would like to thank the personnel and inhabitants of 'De Grevelingen' in 's-Hertogenbosch and of Brunswijck in Eindhoven for their hospitality. Special credit goes to Corrie van Dongen for recruiting participants and her enthusiastic way of teaching seniors how to operate computers.

For leisure activities, lots of fun, and reflections on my work, thanks to the 'Usuals'. Also, I would like to thank my best friends and paranimfen Mattijs Koerhuis and Fabian Super for pep talks and joyful gatherings.

I thank my parents 'Henk and Herma', brother Jan Menno and sister Mariëlle for their support during the years. Lots of thanks go to my girlfriend Joyce. She motivated me to bring my knowledge to the next level and is always a listening ear. Finally I would like to thank our little daughter, Loes Anne, for making life special and more meaningful.

Contents

Chapter 1: General introduction

1.1 Senior computer interaction

Today, people are faced with rapidly changing technologies in their work and home environment, like the Personal Computer (PC) and the Internet. The Internet provides low cost communication in which distances become irrelevant, and communication is enriched by an increased level of social presence (see Bouwhuis, 2000), e.g. MSN Messenger with video chat. The Internet also provides people with the largest information source on the planet. Google has indexed more than 8 billion web pages, and all these pages are available by means of the right query input and a click of the mouse. In western societies, both the PC and the Internet are nowadays affordable, or at least available, to practically everyone. In the Netherlands, in 2006, 88% of the population had access to a PC, and 85% had access to the Internet (CBS, 2006). There are differences between age groups, 57% of people in the age group of 65-75 years had access to a PC, and 50% had access to the Internet (CBS, 2006). Seniors are one of the fastest growing demographics on the web (Nielsen, 2002), and the Dutch SeniorWeb has more than 63.000 members (SeniorWeb, 2006). Not having access or not being able to use current-day technology may put older adults at a disadvantage in terms of their ability to live independently (Czaja & Lee, 2007). Especially for the less mobile seniors, the PC and Internet can be of great help to communicate with others and gather information, but this is only possible if the means provide high accessibility.

Accessibility is strongly related to the concept of usability (Gulliksen & Harker, 2004), as defined by ISO 9241-11 (1998). ISO TS 16071 defines accessibility as: ‘The usability of a product, service, environment or facility by people with the widest range of capabilities’ (see Gulliksen & Harker, 2004). Research in the field of Human Computer Interaction (HCI) and Gerontechnology, focuses on improving the accessibility of digital technologies. HCI is the study of the interaction of humans and computers. HCI is a subdiscipline of computer science (Johnson, 1992), and various subdisciplines of HCI have emerged over the years, which have their grounds in social science disciplines. Gerontechnology is defined as the study of technology and aging for the improvement of the daily functioning of the elderly (Bouma & Graafmans, 1992). Gerontechnology research is focused, among other things, on extending the working phase in life and on maximizing independent years (Harrington & Harrington, 2000). The focus of this thesis is on older adult computer interaction, within the enhancement and satisfaction of communication (Bronswijk, 2006; Harrington & Harrington, 2000), or, more respectfully, without accentuating old, on senior computer interaction. The definition of a senior varies from one country to another. It varies from people of 50, 55, 60, and 65 years of age and older. In this thesis seniors are defined as people from 60 years of age and older; the age most Dutch employees retire. Although all people face difficulties when they interact with computers, seniors have been found to face the most difficulties.

Aging affects general changes (declines) in perception, cognition, and movement control (Bouwhuis, 1992; Charness & Holley, 2004; Czaja & Lee, 2007; Hawthorn, 2000; Mead, Lamson & Rogers, 2002), see table 1.1 for an overview. The effects of aging have an influence on how well seniors use existing technologies as well as how they learn to use new technologies (Charness & Holley, 2004; Morrell et al., 2002; Neerincx, Lindenberg, Rypkema & Van Besouw, 2000; Xie, 2002). Seniors have been found to be slower than young adults on information retrieval tasks (Freudenthal, 2001; Nap, De Greef & Bouwhuis, 2005), on 3D navigation using desktop systems (Sayers, 2004), and on Web navigation (Neerincx et al., 2000). Freudenthal (2001) found that deep menu structures are less suited for seniors than for young adults. In addition, seniors who have been raised in a different technological generation (e.g., the mechanical generation), than young adults, have been found to face difficulties interacting with layered interfaces (Docampo Rama, 2001).

Table 1.1: Age-related changes in perception, cognition, and movement control (Mead, Lamson & Rogers, 2002)

Ability	*Age-related Change*
Vision	
Colour vision	Difficulty discriminating certain wavelengths, particularly blue-greens.
Contrast sensitivity	Increase in minimum luminance contrast needed to resolve high spatial frequency patterns.
Glare sensitivity	Increased susceptibility to glare.
Temporal resolution	Increase in minimum detectable temporal frequency (flash rate).
Visual acuity	Decreased ability to resolve small details.
Visual selection	Difficulty selecting relevant information in a display that contains relevant and irrelevant information.
Audition	
Auditory frequency range	Decline in sensitivity to high frequencies.
Auditory sensitivity	Decreased ability to distinguish among tones.
Auditory selection	Decreased ability to separate speech from background noise.
Cognition	
Working memory	Reduced capacity to maintain information in active memory.
Spatial visualization	Reduced ability to hold and operate on spatial representations in working memory.
Language comprehension	Decreased ability to process complex text.
Episodic memory	Poorer explicit memory for specific events and their contexts. Slower acquisition of new knowledge.
Semantic and procedural memory	Previously acquired general knowledge and skills are well maintained. Slower acquisition of new skills.
Movement control	
Fine motor control	Decreased ability to manipulate very small controls.
Noise to force ratio	Decline in accuracy of rapid movements.

Aging is associated with longer deliberation times, poorer decision making, and reduced risk taking, demonstrated by performance on a computer based gambling task (Deakin, Aitken, Robbins & Sahakian, 2004). Ellis and Allaire (1999) found that there is a negative correlation between age, computer knowledge, and computer interest. Although seniors have attained a rich technology knowledge base due to many years of experience in the home- and work environments, most seniors have acquired little knowledge of recent technologies such as the computer, e.g. many are still inexperienced with the Internet (Brynin, Raban & Soffer, 2004). Even if their experience is similar to their younger counterparts, older novice users are less successful in operating a library retrieval system (Mead, Sit, Rogers, Jamieson & Rousseau, 2000) or in navigating web sites (Mead, Batsakes, Fisk & Mykityshyn, 1999).

Despite these cognitive declines, people at all ages can learn to use new communication and information technologies, although it takes people more time to learn when they get older (Hawthorn, 2000; Kelley & Charness, 1995). But, are older people willing to spend this time in learning to use new technologies? Morrell, Mayhorn and Bennett (2000) found that old-old adults (ages 75-92) have the least interest in using the web compared to middle aged adults (ages 40-59) and young-old adults (ages 60-74). Seniors perceived the benefits of technology and commercial products as more important than the costs (Melenhorst, 2002; Sharit, Czaja, Perdomo & Lee, 2004). Melenhorst (2002) has indicated three major preconditions for the adoption of new communication technology by seniors. First, the potential benefits of a new medium should be relevant from the perspective of older adult users, with respect to their specific communicative aspirations. Second, these relevant benefits should be explicit and clear, and third, the potential costs involved in using a new medium should be transparent to the older adult user, especially when the user is inexperienced, in order to reduce uncertainty about the attainment of the benefit. There are differences between young adults and seniors in learning and using new technologies, but is it necessary for the accessibility of digital information systems to make a distinction on the base of age? In addition, how could systems be designed to overcome age-related problems?

Of course, not all seniors face difficulties when interacting with computers; the older a segment of the population is, the more diverse it will be in terms of capabilities and limitations (Czaja & Lee, 2007; Ellis, 2005). Most seniors are very well capable of acquiring computer skills, learning how to use the web (see Morrell et al., 2002), or learning how to use PDA's (Mayhorn, Lanzolla, Wogalter & Watson, 2005). The majority of older people nowadays, and increasingly so in the future, do not fit the stereotype of people with a loss of some sort (Bouwhuis, 2003). The larger part consists of mentally and physically healthy autonomous adults between 55 and 75 years of age (Bouwhuis, 2003). According to Dickinson and Dewsbury (2006), it is important to avoid treating ageing in terms of functionality decline and incapacity, because it is unlikely to be recognized as a fair or accurate representation of the life experience by most adults over 65. Newell (2006) states that 'high functioning' older people have similar characteristics to 'medium functioning' middle aged people. Thus, in terms of their abilities, design that is appropriate for older people will be appropriate for most of the population, whereas design for younger and middle aged people can exclude significant numbers of older people (Newell, 2006). Ellis (2005) argues that the older population should not be treated as a special population, inclusive research should be conducted that takes older adults into consideration as part of the 'normal' population to the greatest extent possible. Still, there is a considerable group of seniors that do have problems interacting with computers due to very specific age related changes or technology generation specific knowledge. A seemingly more widely accepted approach to treat the older population as a special population is the development of 'Personas' (Newell, 2006; Wickens, Lee, Liu & Gordon-Becker, 2004). Personas are not real people, but they represent key characteristics of a user population in the design process (Wickens et al., 2004).

There are several constraints that should be considered in the design of accessible ‘Information and Communication Technology’ (ICT) for seniors; age related declines in cognition, perception, and motor control, the level of experience, longer learning times, and the perceived benefits. To provide high accessibility, experimental evaluations and iterative design with representative users are essential (see Wickens et al., 2004); they provide feedback for making modifications to the design, and the evaluations shed light on the specific support that is needed to overcome the above constraints. As mentioned, the older a segment of the population is, the more diverse it will be in capabilities and limitations. To provide accessibility of a device for such a diverse user group, personalisation at the content and interaction level by means of user profiles could be used (see Cremers & Neerincx, 2004). The user profile can contain specific characteristics of the user, for example personal details (e.g. name, gender, age), capabilities (e.g. memory, sight), and personal interests. The information in the user profile can be used by the device to adjust the way content is presented and the interaction modalities and styles that are optimally suited for the individual user. Guidelines are also important to consider in the design of accessible ICT for seniors. General guidelines for interface design (see Nielsen, 1994; Shneiderman, 1997) also apply to seniors. However, specific guidelines for seniors have also been developed that account for the changes in perception, cognition, movement control and system knowledge, see table 1.2 for an overview. According to Wickens et al. (2004) many guidelines are just that: guides rather than hard and fast rules. Interface design guidelines are helpful in the development of a first prototype, but do not provide a template of the interface design. Guidelines do not provide the type of domain information (content) that is appropriate for the user, how this information should be presented, and what technologies can be used to increase accessibility to digital information.

Table 1.2: Interface design guidelines for seniors (Echt, 2002)

Typography

Guideline	*Improves*
Strive* for 14 pt. font size for body text and 18-24 for headers	Reading Acuity, Reading Speed, Resolution
Use sans serif font types like Geneva or Helvetica	Reading Acuity, Reading Speed
Use bold; Avoid italics	Search
Use mixed case; avoid using uppercase text	Lateral Masking, Reading Speed
Use left justified text; avoid centered or full justification	Reading Speed, Eye Movements
Strive* for leading greater than single spacing	Reading Speed, Lateral Masking
Line Lengths between 45-60 characters	Eye Movements, Search, Reading Speed

Layout

Guideline	*Improves*
Use black text on white ground	Acuity, Contrast Sensitivity, Glare
Use white space actively	Search, Reading Speed
Avoid multi-column format or frames	Eye Movements, Distraction
Use consistent placement of page elements	Search

Organization

Guideline	*Improves*
Reduce text from print by 50%	Search, Legibility, Distraction
Minimize number of hypertext links in a line of text	Search, Distraction
Separate steps of a procedure using #s or bullets	Search
Use headings or subheadings	Search, Organization
Present an alphabetical index	Search
Cue important information	Search, Eye movements

Navigation

Guideline	*Improves*
Use large buttons with symbols and text	Search, Mouse Coordination
Provide visual feedback upon selection	Mouse Coordination
Minimize need for scrolling	Search
Increase the size of peripheral elements	Peripheral Detection
Use yellows/blues for important information in the periphery	Peripheral Detection

Graphics

Guideline	*Improves*
Strive for consistency, simplicity, and meaning	Search, Distraction
Strive for clear graphics	Contrast Sensitivity
Place text on unpatterned backgrounds	Acuity, Contrast Sensitivity, Search
Use colour judiciously, consistently & with purpose	Search, Selective Attention
Avoid using similar hues, pastels, or blues/greens together	Colour Discrimination
Avoid flashing/blinking text	Distraction

Continued

Other

Guideline	*Improves*
Be aware of how information looks printed	Legibility
Allow distance and height of monitors to be adjustable	Presbyopia, Glare
Encourage the use of high resolution screens	Reading Acuity, Contrast Sensitivity, Reading Speed, Search, etc.

* = font size and line lengths should be adjustable by the user.

1.2 Information retrieval

As was discussed in the previous sections, there are problems with accessibility for seniors. Specific design guidelines for seniors have been developed that can be used to overcome problems they face when they interact with computers and the Internet. But, guidelines do not provide a template of what content is appropriate for the user, how this content should be presented and what technologies could be used to increase accessibility to the content. Possible answers to these questions can be found within the field of ‘Information Retrieval’ (IR).

At an applied level, information retrieval likely has his roots in the ancient Greek and Roman libraries. The thousands of scrolls should have been systematically documented in order to effectively and efficiently retrieve the right scroll when searched for. Information science is a young discipline, the earliest formal use of the term information science dates back to 1958 (Ingwersen, 1992). With the development of the personal computer, the computer sciences faced similar difficulties with information retrieval as information science studies. Just like the Greek librarian and his scrolls, computer scientists need to carefully think where to store a bit of data, so it can be effectively and efficiently retrieved. At the database and interface level, similar considerations need to be applied. Information retrieval systems emerged in the 1950's and 1960's as static, batch processing systems, in the 1970's the access to IR systems became dynamic and interactive (Saracevic, 1996).

When people engage in information-seeking behaviour, it is usually because they are hoping to resolve some problem, or achieve some goal, for which their current state of knowledge is inadequate (Belkin, 2000). According to Saracevic (1996), research on interactive aspects of IR has not reached maturity; it may be even said that it is barely emerging out of infancy. Anderson & Pérez-Carballo (2001) focused on humans (users) in information retrieval research. The effectiveness of an IR system is highly dependent on the input of the user.

The way information is retrieved from a computerized IR system has been used as an analogy for the way humans retrieve information from memory. IR can be seen as occurring in three stages. First, retrieval cues are generated, second, the cues are used to retrieve information, and third the retrieved information is verified (see Foltz, 1991). This is an oversimplified way of describing how human memory works. There are differences between human memory and computers. Human memory is context dependent, human memory excels in capacity and flexibility, and humans forget (Baddeley, 1997). There are also differences in the way humans and computers index information. According to Anderson and Pérez-Carballo (2001), machines are very far from simulating the work of a human indexer. Human indexers interpret the text, in the context of their cultures and their personal experiences, including their prejudices, taking into consideration user needs and desires (Anderson & Pérez-Carballo, 2001). Computers have distinct capabilities that humans lack; they can index and retrieve a huge amount of information in a short time. For example, the search string "Information Retrieval," in Google, provides the user with almost 3 million pages within a second (Google, 2007).

When human users employ computers for information retrieval they have to submit search terms to an IR system, which performs some type of matching and then presents a list of retrieved items for users to evaluate (Anderson & Pérez-Carballo, 2001). As an example, Swain and Ballard (1991) introduced a simple automated technique to index (retrieve) similar-looking images from an image database. The colours of an image are indexed into a colour histogram, and by vector analysis, similar looking images can be retrieved. To automatically index or retrieve text, Latent Semantic Indexing (LSI) or Latent Semantic Analysis (LSA) can be used. The technique is based on a word by document matrix. Words, sentences, or whole texts represent points in a multidimensional space, and the 'semantic' relation between them can be extracted (see Hoenkamp, 2003; Kintsch, 1998; Landauer, Foltz and Laham, 1998). The technique is similar to the more common factor analysis, and today the technique is used, for example, to compare student papers for authenticity and to give students feedback about the content of their writing.

As mentioned, both humans and machines have distinct capabilities in information retrieval. It depends on the information domain what tasks and functions are preferably allocated to a person or the technology, and it depends on humans' choice to be in control. In 1951, Fitts introduced the discussion of delegation of control, or the allocation of tasks and functions to a person or technology with the MABA-MABA concept. MABA-MABA stands for 'Men Are Better At – Machines Are Better At' (Fitts, 1951). Humans are very capable of narrowing their search information by choosing meta data, and meta data variables, and with these variables, a machine can rapidly retrieve the information from a large information database.

To provide a machine with information from a domain, a data model (i.e. ontology or conceptual model) needs to be developed. First, the current situation should be examined. How are the concepts within a domain stored in books or other information systems? What concepts are used in these information systems? In addition, what structure could be optimal for retrieving the information? Experts in the field can comment upon the data model that is created from the current situation. After that, a first prototype can be evaluated with future users by means of low-fidelity prototyping. Do these users understand what the concepts mean? Do they use other concepts for the same information? Can they effectively, efficiently and satisfactorily retrieve the information? Iterative design, paper prototyping (see Rettig, 1994; Sefelin, Tscheligi & Giller, 2003; Virzi, Sokolov & Karis, 1996), and experimental evaluations with representative users can be used to find answers to these questions.

1.3 NARRATOR – A health care related information retrieval system

In a world in which the amount of information and the retrieval rate is continuously rising, the demand for health care information retrieval systems is also increasing. Ever more seniors are getting online to search for health related information; in the United States of America there was a growth rate from the year 2000 to 2003 of 25% (Fox, 2004). The research presented in this thesis was part of the NARRATOR project and funded by the Netherlands Organisation for Scientific Research (NWO). NARRATOR is the name of the project for Narrative Disclosure of Health Care Knowledge. The goal of the project was to gather functional and non-functional requirements for an information retrieval system that provides breast cancer patients[1] with narratives of other patients and ways to communicate with each other. With the gathered requirements, a working prototype can be engineered and evaluated with future users. Degner and Sloan (1992) suggested that the psychological impact of cancer may influence most patients to prefer a passive role, at least until they had have an opportunity to learn more about their disease and its treatment. The user group that NARRATOR focused on were breast cancer patients that have already had treatment and are searching on the Internet for narratives of other patients and ways to communicate with each other.

The features of current websites for breast cancer patients are insufficient to retrieve breast cancer patients' illness stories (Overberg, Toussaint & Zwetsloot-Schonk, 2006). Overberg, Alpay, Verhoef & Zwetsloot-Schonk (2007) studied the preferences of breast cancer patients regarding content, appearance, and search options, and found that patients are mainly interested in fellow patients' experiences about how to cope with emotions, the impact of cancer in daily life, and physical discomforts. Regarding the appearance, most patients prefer a section of an illness story, and some want to be able to click on the corresponding complete story. Regarding the search options, the majority of patients wants to be able to select illness stories on the basis of several authors' features; i.e. treatment undergone, age, presence of metastases, time since diagnosis, and whether or not they are caring for children.

The diagnosis of breast cancer is traumatic for patients and their family. Post Traumatic Stress Disorder symptoms have been found to be prevalent among women with breast cancer (Levine, Eckhardt & Targ, 2005), or at least patients show intense negative emotional reactions to breast cancer (Palmer, Kagee, Coyne & DeMichele, 2004). According to Case, Andrews, Johnson & Allard (2005), people

[1] Initially, NARRATOR started with the focus on falling incidents in elderly people.

sometimes avoid information, particularly in the context of health care. Avoiding information is closely linked to feelings of anxiety and fear, and if people pay attention to the information, it will cause mental discomfort or dissonance (Case et al., 2005). Because of this, it is important that an information system that provides breast cancer patients with stories of other people with breast cancer is designed in such a way that it does not bring about additional stress and negative feelings.

The user group under study in this thesis consists of seniors, because knowledge about senior computer interaction also includes knowledge about a large group of breast cancer patients (mean age of diagnosis in the Netherlands is 60 years (RIVM, 2007), a group that will most likely face difficulties interacting with a computer.

1.4 Stress in senior computer interaction

In this subchapter, a historical background of stress is presented first to illustrate the concept of stress and the associated research that has been executed. After this introduction, the literature about stress in seniors is discussed along with differences in stress among people from different age groups. The small amount of research that has been carried out to understand stress in senior computer interaction is presented, and common methods to measure stress are discussed.

1.4.1 History of stress

Stress must have occurred even to prehistoric man, although he may not have been consciously aware of it. When the feeling came he must have realized instinctively that he had exceeded the limits of what he could reasonably handle (Selye, 1970). In the seventeenth century, the concept of stress was introduced in mechanical engineering. At the end of the seventeenth century, Robert Hooke published the Lectiones Cutlerianae, which contains among other discoveries Hooke's Law: stress is proportional to strain (Andrade, 1960). It was also in the seventeenth century that according to Descartes the body of an animal was nothing more than a complex machine and it was thought that the human body did not differ essentially from mechanical puppets (Boon, 1996). The mechanics of a machine were used as a model to understand the human body, and the stress and strain of machines and constructions were also to be expected to be apparent in humans.

In the eighteenth and nineteenth century, the term *"neurasthenia"* came into being to describe a diagnostic syndrome that referred to weakness or exhaustion of the nervous system (Berger, 1973). Sigmund Freud described it as a "fundamental disorder in mental functioning" (AllRefer, 2007), and used cocaine as a treatment for neurasthenia (Clifford & Scott, 1956). Beard, an American psychiatrist, who also worked with Edison, used the analogy of an electrical system: when the dynamo was overloaded, the lights of the periphery dimmed. Beard regarded "neurasthenia" as a strictly nineteenth century condition related to technical advance (Berger, 1973). Although neurasthenia is hardly used as a diagnosis anymore, the nineteenth-century notion that the stresses of (modern) life could cause wear and tear of the nervous system is a notion that found much scientific ground into the twenty-first century.

In the nineteenth century, Claude Bernard described the *"milieu intérieur"* of animals that was characterized by constancy and balance of the internal environment (Dohms & Metz, 1991; Robin & Bromberg, 1959). Many bodily conditions must not only remain constant, they must stay the same, or at least remain within rigidly set limits (McEwen & Lasley, 2002). Other references to self-regulatory arrangements are by Pflüger (1877), Léon Fredericq (1885), Charles Richet (1900), and even Hippocrates (460-377 B.C.) (Cannon, 1932). Claude Bernard laid the groundwork for the scientific study of stress, but the first scientist to actually bring the word into the medical vocabulary was a physiologist by the name of Walter Cannon (McEwen & Lasley, 2002).

In the beginning of the twentieth century, Cannon laid the groundwork for the introduction of the psychosomatic approach (Cooper & Dewe, 2004). Instead of describing and understanding humans solely by his mechanical parts and the consequences of malfunction for disease, thoughts, motives, and societal influences were considered as important causes for disease. Cannon introduced the concept of *"homeostasis,"* which is similar to the *"milieu intérieur"* concept of Bernard. Homeostasis comes from the Greek roots *homeo*, meaning same, and *stasis*, meaning stable – remaining stable by staying the same (McEwen & Lasley, 2002). The internal balance of an organism is maintained by constantly adjusting internal (set) values in concurrence with external (actual) values. Cannon used the word stress in terms of the emotions (e.g. great emotional stress), and he also ushered in the more "engineering" concept, discussing the strain that cold, lack of oxygen, or low blood sugar level can place on the body (McEwen & Lasley, 2002). Cannon also introduced the "fight or flight" response to fear and anger. Fear is associated with the instinct to run, to escape; and anger or aggressive feeling, with the instinct to attack (Cannon, 1932). During these emotions, respiration deepens, the heart beats more rapidly, the arterial pressure rises, the blood is shifted away from the stomach, adrenalin is released from the adrenal medulla, and other transformations occur in the body (Cannon, 1932).

In 1908, Yerkes & Dodson published a study about the relation between stimulus strength and habit formation for tasks varying in discrimination difficultness (Teigen, 1994). In 1955, Hebb introduced his own version of their results, which most scientists know today as the Yerkes-Dodson law. In the Hebbian version of the Yerkes-Dodson law, an inverted U describes the relation between arousal and performance (see Teigen, 1994). The Hebbian inverted U-shaped curve demonstrates that performance is dependent on arousal, too little or too much arousal will decrease performance, and optimum performance lies in the middle. The Hebbian version states that high levels of stress, anxiety, or motivation produce a monolithic impairment of performance (Diamond, Campbell, Park, Halonen & Zoladz, 2007). While the original version based on the actual Yerkes-Dodson findings takes into account the finding that strong emotionality can enhance performance under "simple" learning conditions, and impairs performance under more complex or challenging learning situations (Diamond et al., 2007). The Yerkes-Dodson Law arousal concept has been interchanged with e.g. motivation, emotionality, tension, anxiety, stress, and even workload (see Teigen, 1994), as if the choice of concept does not make a difference.

In the 1930s, Hans Selye began research that put the word *stress* on the map, ultimately reintroducing it into the English language (McEwen & Lasley, 2002). Hans Selye introduced the theory of a 'General Adaptation Syndrome' (GAS), which is the body's generalized response to challenges. GAS consists of three stages; the first stage is the 'alarm reaction;' the second stage is 'resistance,' and the third, when stress was persistent, 'exhaustion' (Doublet, 2000). After much opposition from the scientific community, he added the concept 'stress' to his theory (Doublet, 2000). It is interesting to note that according to Selye, stress is not a deviation from homeostasis, because any biologic function causes marked deviations from the normal resting state in the active organs. Furthermore, stress it is not anything that causes an alarm reaction; it is the stressor that does that (see Doublet, 2000).

Until the 1960's stress research was conducted mainly in the field of physiology, and little attention was paid to psychological stress. Not everyone reacts the same way to the same events, and this lead to theories involving the concept of *"appraisal"* (see Doublet, 2000). According to Lazarus and DeLongis (1983), appraisal refers to the way a person construes the significance of an encounter for his or her well-being, that is, as irrelevant, benign, harmful, threatening, or challenging, the latter three being forms of stress appraisal. Coping processes are initiated to manage a stressful encounter (see Cooper & Dewe, 2004; Lazarus & DeLongis, 1983), and how a person copes with this encounter determines his or her emotional response (Lazarus & DeLongis, 1983). The two main classes of coping strategies are problem-focused coping and emotion-focused coping (Lazarus & Lazarus, 2006). In problem-focused coping, a person's attention centers on what can be done to change the situation to eliminate or lessen the stress, whereas in emotion-focused coping, no effort is made to change the situation but rather the way a stressful encounter is construed or attended to and therefore the emotional reaction to it (Lazarus & DeLongis, 1983; Lazarus & Lazarus, 2006).

Stress is an English concept and it took a while before the concept was included in the Dutch dictionary. The concept of *"stress"* appeared in the Dutch Van Dale dictionary after 1950 (Van Dale, 2003). In the tenth edition (Kruyskamp, 1976) the meaning of stress was: "in de geneesk. aanduiding van een toestand die bepaalde afweermechanismen in werking doet komen." Translated into English this means: "in medicine, reference to a state that activates certain defence mechanisms." In the eleventh edition (Geerts, Heestermans & Kruyskamp, 1984), stress was defined as: "toestand, waarin het evenwicht van de bio-fysiologische functies in het lichaam is verstoord door grote lichamelijk of geestelijke spanning en die bepaalde afweermechanismen in werking doet komen". Translated into English this means: "state, in which the balance of bio physiological functions in the body is disrupted by large physical or mental strain, and which activates certain defence mechanisms”. It took at least fifty years before the concept of stress, as defined by Cannon and Selye, had a proper definition in the major dictionary of contemporary Dutch. Nowadays the concept of stress is common in verbal communication, e.g. students are stressed out over studying, parents are stressed out over newborns, and employees are stressed out by work.

At the end of the 1970’s Bandura (1977) and Ursin, Baade and Levine (1978) laid the groundwork of their current theories of stress. Bandura (1977) found support for the view that perceived self-efficacy mediates anxiety arousal. According to Bandura (1977), personal efficacy determines whether coping behaviour will be initiated, how much effort will be expended, and how long it will be sustained in the face of obstacles and aversive experiences. Ursin et al. (1978) found that ‘coping,’ the trust in one’s own abilities to perform, influences subjectively reported fear and the stress responses. They (Ursin et al., 1978) expected that arousal would be reduced when subjects’ performance would be at an acceptable level; it turned out that positive outcome expectancies reduced arousal, independent of performance. It is difficult to explain the difference between Bandura’s self-efficacy concept and Ursin’s coping concept. H. Ursin (personal communication, October 25, 2006) and Bandura appeared to agree that Bandura is talking about something here and now, like the self-efficacy of an athlete before one specific competition event, the coping concept appears to generalize from one response to all or most.

In the 1980's a new term came into being within the field of stress: *"allostasis"*. It comes from the Greek root *allo*, meaning variable, and *stasis*, meaning stable (Ropeik, 2004). Allostasis emphasizes the point that allostatic systems help keep the body stable by being themselves able to change and to provide enough energy to cope with any challenge – not just the life-threatening ones (McEwen & Lasley, 2002). Allostasis is similar to Bernard’s “*milieu intérieur*”, and Cannons' *"homeostasis"* and *"fight or flight"* response, although allostasis is more in line with natural selection and the organism’s ability to adapt to his environment. If people only try to remain stable by staying the same (homeostasis), they cannot respond to abrupt changes in the environment (see Sterling, 2004). In situations in which people cannot cope with a challenge, allostatic load denotes a failure of the body's efforts to remain stable. According to McEwen & Lasley (2002), allostatic load is like two sumo wrestlers on a seesaw – the seesaw may be in balance, but it is under a strain that may eventually cause it to break. Allostatic load refers to the ‘wear and tear’ that the body experiences due to repeated cycles of allostasis, and it can be created by for example unhealthy diets, low social support (loneliness), and aging (see McEwen & Lasley, 2002).

In 2004, Ursin and Eriksen (2004) published the Cognitive Activation Theory of Stress (CATS), which is a result of a long series of experiments and theoretical papers, with data from animals and humans (Eriksen, Murison, Pensgaard & Ursin, 2005). CATS offers an insight into the psychological mechanisms explaining when a stress (alarm) response occurs and when it may become maladaptive (Ursin & Eriksen, 2004). The basic rules for when stress occurs are the same across cultures and species, from fish to Olympic performance in humans (Eriksen et al., 2005). Four aspects of stress are considered in CATS: the stimuli (stressors), the stress experience, the stress response, and the experience of the stress response. According to Ursin and Eriksen (2004), all four aspects of stress can be measured. Whether a stimulus is pleasant or threatening depends on the individual appraisal of the situation, which is based on previous experience and expectations of the outcome. Humans report a threatening or negative perceived stimulus as stress, i.e. a stress experience. The stress response is a general, unspecific alarm response occurring whenever there is a discrepancy between a set value (what is expected or the normal value), and an actual value (what is happening in reality), i.e. an imbalance in allostasis. The perceived probability of eliminating such discrepancies influences access to the arousal system and the probability of success has consequences for the hierarchy of set values (Ursin & Eriksen, 2004). The feedback from the stress response is stored as response outcome expectancies, these can be positive, negative or no expectancy.

From the work of Cannon, Selye, Lazarus, Bandura, Ursin, Baade and Levine, McEwen, and many others, stress had become a major research subject within the discipline of immunology, physiology and psychology.

1.4.2 Stress in seniors

When discussing stress in seniors the obvious question arises whether there would be any difference between age groups. On a psychological, physiological, and immunological level, differences have been found between age groups on stress experiences and responses.

In the nineteen-sixties, Holmes and Rahe (1967) developed ‘The Social Readjustment Rating Scale’ (SRRS) that scores major life events and the impact it has on illness. Every major life event is assigned a score from 100 for death of spouse, to 11 for minor violations of the law. Forty-three life events are used in the SRRS and the score of the testee is the sum of the points assigned to each life event that is applicable. A score higher than 300 represents an extremely highly stressful life. It is apparent that the chance of the occurrence of the loss of relatives increases with aging, and that would consequently lead to seniors scoring high on the Holmes Rahe scale. The problem with the Holmes Rahe scale is that people differ in reactions to major life events, as discussed before about the appraisal concept. Old age has been viewed as a time when stress factors, such as health decline and retirement, are viewed as challenges the individual’s sense of control (Karel, 1997). However, adult developmental research has found more evidence for a stable sense of control across adulthood, although achieved through changing mechanisms (Karel, 1997). Stressful events appear to be most frequent in the lives of young adults, who report high numbers of life stress events, and very old adults, many of whom struggle with chronic illness and disability (Karel, 1997). Lazarus and DeLongis (1983) found no satisfactory answer to the question of how the dynamics of stress and coping change with the circumstances of living and the processes of aging. Personal beliefs, values and commitments, which develop from a person's unique history, shape the appraisal of stress and the manner in which stress is coped with (Lazarus & DeLongis, 1983).

The aging brain becomes more vulnerable to the effects of stress (McEwen & Lasley, 2002). Glucocorticoids are essential for surviving acute physical stress but they may cause adverse effects when secretion is sustained (Sapolsky, 1996). With increasing age, rats tend to lose a down-regulation in glucocorticoid receptors when exposed to repeated stress, which gives stress-generated cortisol a broader target area (McEwen & Lasley, 2002). Cortisol levels increase with aging (Van Cauter, Leproult & Kupfer, 1996), and human elderly subjects show cognitive impairments that seem to be related to cortisol levels (McEwen & Lasley, 2002). According to McEwen and Lasley (2002), older people may be uniquely vulnerable to stress, because of the intricate connections between allostatic load and brain aging.

The immune system protects an organism against diseases. Aging and stress have a negative effect on the functioning of the immune system (see Hawkley & Cacioppo, 2004). Stress results in the secretion of many hormones that can have profound effects on bodily processes and vital organs, including the immune system (Lazarus & Lazarus, 2006). Research on stress in older adults provides evidence that increased stress exposure, inadequate stress buffering, exaggerated stress reactivity, extended stress duration, and possibly diminished restorative processes have effects that mimic, exacerbate, and sometimes accelerate the effects of aging on immunity (Hawkley & Cacioppo, 2004).

In conclusion, it appears that seniors are more vulnerable to stress than young adults. The engineering analogy of stress and allostatic load could clarify why. Older structures are less able to withstand heavy loads (stress), because of many years of repetitive strain, i.e. 'wear and tear'. Health problems, death of relatives and limited control over life contribute to this strain. On the other hand over the years seniors have found ways to cope with stressors. This may be an important factor why seniors might be stressed interacting with new media, because time to find ways to cope with new stressors is limited and costly.

1.4.3 Stress in senior computer interaction

March 11, 2004
Young gets Old on the Internet

“Yesterday, ten seniors, students of a five week Internet course at the Prisma College Graaf Engelbrecht, received the certificate (passed the internet exam).”

“Some course participants did not show up for the exam, although they properly excused, with or without the use of e-mail. Cause for excuse: exam fear.”

Figure 1.1: Example of exam fear in senior computer interaction, translated from Dutch into English (Krijnen, 2004)

The above example shows that there may well be more to senior computer interaction than an accessible interface, system knowledge, and motivation, if seniors excuse themselves from a computer course exam with the reason: exam fear. Within the computer domain, seniors are less confident in their judgments than young adults, and this is related to poorer, computer-related, global self-efficacy beliefs (Marquié, Jourdan-Boddaert & Huet, 2002). The study of Marquié et al. (2002) suggests that low confidence in their ability is one reason why older people have difficulty in mastering new computer technologies. Hawthorn (2003) noticed that many age restricted users appear to operate with a low expectation of success; they were frequently uncertain about their actions (e.g. sending e-mail) and asked if it was successful. Czaja, Charness, Fisk, Hertzog, Nair, Rogers, and Sharit (2006) suggested that it may be helpful to provide older people with some type of stress inoculation training that teaches them to better deal with their performance anxiety. Czaja et al. (2006) found that computer self-efficacy is an important predictor of computer anxiety, and because of this it is important that older people receive encouraging feedback during training and experience some level of success.

There must be some relation between fear of failure, anxiety or stress, and senior computer interaction. Computer-based work has been found to be related to stress in seniors (Birdi & Zapf, 1997; Czaja & Sharit, 1993), although Birdi and Zapf (1997) have not studied this experimentally and Czaja and Sharit (1993) focused primarily on mental workload. Furthermore, there are differences between males and females. Female seniors have been found to report more anxiety and less computer knowledge than male seniors (Karavidas, Lim & Katsikas, 2005). There is a lack of evidence about the direct impact of software design on stress (Hamborg & Grief, 2003) and Hawthorn (2000) suggested an exploration of how stress affects, and is affected by, interface use for older users.

If seniors use computers, it helps them to increase self-efficacy and lower computer anxiety, thereby increasing overall life satisfaction (Karavidas et al., 2005). The implicit benefits may include, among other things, allowing the older adult users to be more independent, maintaining a social network of friends and families, and staying informed about the health social network of friends and families (Karavidas et al., 2005). Slegers (2006) found that the level of computer anxiety of inexperienced seniors did not decrease as a result of using a computer and the Internet. Because more extensive use of computers does not result in decreases in computer anxiety, Slegers (2006) suggested that other solutions should be focused on such as removing psychological barriers seniors experience (psycho education).

The challenge is to decrease stress and to increase confidence in senior computer interaction. Computer use could help to increase self-efficacy, but low confidence in ability to perform hinders computer technology mastery, and may even cause stress. There should be a way, aside from computer usage, that increases self-efficacy and decreases stress in seniors. The aim of technology should be to facilitate, to be an aid and not an obstacle (Marcellini, Mollenkopf, Spazzafumo & Ruoppila, 2000).

1.4.4 Stress measurements

Many different psychological and physiological stress reactions have been found that provide a measure of stress (Salas, Driskell & Hughes, 1996). When humans perceive a set of stimuli as threatening or negative, then they report this as stress, and this stress experience is perhaps the most relevant in human stress research in working life (Ursin & Eriksen, 2004). The stress experience can be measured subjectively by questionnaires measuring e.g. fear, anxiety and annoyance, for example the stress scale of Lundberg and Frankenhaeuser (1980), or the Dundee Stress State Questionnaire (see Matthews, Campbell, Falconer, Joyner, Huggins, Gilliland, Grier & Warm, 2002). The stress scale of Lundberg and Frankenhaeuser (1980) consists of six concepts (3 positive and 3 negative stress reactions) that have to be rated on a scale. The scale takes little time to administer, has been found to correlate with physiological responses, and is still used today (Hjortskov, Rissén, Blangsted, Fallentin, Lundberg & Søgaard, 2004). The Dundee Stress State Questionnaire is lengthy; it contains of eight pages and takes about 15 minutes to administer.

Stress can be also measured physiologically the pulse rate, heart rate (HR), heart rate variability (HRV), systolic and diastolic blood pressure (BP), glucocorticoid (e.g. cortisol) output and palmar sweating (i.e. skin conductivity). HR is the number of ventricular contractions per minute. Pulse rate is the frequency of pressure waves in the arteries and is the same as HR in normal, healthy individuals. BP is the internal pressure in the arteries near the heart (Kroemer, Kroemer & Kroemer-Elbert, 1997). Systolic BP refers to the pressure of blood when the heart contracts and diastolic BP refers to the pressure of the blood when the heart relaxes. The stress reaction is coupled with an increase in HR, a decrease in HRV, an increase in BP, higher levels of cortisol secretion (which increases BP), and an increase in skin conduction. When people exercise, HR and BP also increase; furthermore when people are concentrated performing a task HRV decreases. Cortisol levels are higher in the morning, compared to in the evening, and rise in reaction to meal times (Lovallo & Thomas, 2000). Skin conductivity also increases when people are emotionally aroused, for example when they think about a handsome man or pretty woman. Only the 'stressors' that contain an element of frustration, anger, anxiety and any other excited states, increase BP (Doublet, 2000).

Self-report measures have been found to be more reliable than some of the physiological measures, HR and BP (Shostak & Peterson, 1990). According to Doublet (2000), self-reports of feeling more or less 'stressed' are plagued with the same problems inherent to all self-reports. Instead of feeling, for example 'annoyed', 'worried', or 'frustrated', people seem more likely to declare that they are feeling 'stressed' even though it does not necessarily accurately describe to themselves and other what they feel (Doublet, 2000).

1.4.5 Definition of stress

The use of the concept of stress is still gaining in popularity by lay people and scientists, but a widely accepted theory and definition is still lacking. Some scientists even question the existence of stress (for example Doublet, 2000). Although the discussion about the existence of stress is relevant for stress related research, it falls out of the scope of this thesis. For the presented research that follows the focus is on the stress people experience and report in their daily lives, and specifically to the subject addressed, if seniors experience stress in computer interaction, and if so, why they experience stress and how their stress experience can be reduced. In this thesis, the definition of stress is based on the Cognitive Activiation Theory of Stress (CATS) (Ursin & Eriksen, 2004). Stress is defined as a physiological response to, and psychological experience of, an imbalance between what is expected and what is happening in reality. When the expected event is perceived as negative, then there likely will be an increase in physiological- and psychological stress.

1.5 Research problem and outline of the thesis

So far, no experimental data are available for a (causal) relation between the accessibility of digital information and stress in seniors. In addition, little is known about why seniors are stressed during computer interaction, i.e. what mechanisms contribute to stress in senior computer interaction. The goal of this research is to gather knowledge about the factors that influence stress in senior computer interaction, within the information retrieval domain.

Experiments are conducted in which seniors perform search tasks on retrieval systems. As mentioned, the research presented in this thesis is part of the NARRATOR project with the goal to gather functional and non-functional requirements for an information retrieval system that provides breast cancer patients with narratives of other patients and ways to communicate with each other. To gather these requirements, we started out within the art domain with non-patients not to inordinately hinder patients (see also Blanson Henkemans, Rogers, Fisk, Neerincx, Lindenberg & Van der Mast, 2007). The art domain was chosen because it is recognizable and pleasurable for many people; hence it is unlikely to elicit negative feelings that could influence interaction performance. In the first experiments, an image retrieval system is iteratively engineered and evaluated to provide high accessibility for all. Information is gathered from behaviour in tasks, about accessibility of different designs. For breast cancer patients it can be emotionally demanding to search for narratives of other patients and information about their disease. For this, it is important that they do not experience additional stress and negative feelings during information retrieval caused by complex systems. In the subsequent experiments presented in this thesis the effect of complexity on acute stress is measured, and the effect of seniors' expected task performance on acute stress.

1.5.1 Research outline

Chapter 2 addresses the design and experimental evaluation of an accessible retrieval system for seniors. In this chapter, a method to design a retrieval system for seniors is discussed. Cognitive engineering principles, iterative paper prototyping and experimental evaluations were used to engineer a high fidelity working prototype with which art images can be retrieved. Three evaluation experiments and an error analysis were performed to gather knowledge about interface requirements for an accessible retrieval system for seniors.

Chapter 3 examines the influence of complexity on acute stress in senior computer interaction. There is a lack of evidence about the influence of interface design on stress in seniors. In an experiment, low- and high complex art image retrieval systems were compared for their effect on acute stress (physiological and subjective) in seniors, when they perform search tasks on the two systems. It was expected that acute stress would be higher when seniors perform search tasks on a highly complex system than on a system of low complexity.

Chapter 4 examines the effect of complexity and induced expected task performance on acute stress. A theoretical model of stress in senior computer interaction was developed. Two factors could influence acute stress in senior computer interaction. The first factor is the complexity of a system and the second factor is expected task performance. It was expected that acute stress would be higher when seniors perform search tasks on a high complexity system than on a low complexity system. In an experiment, acute stress was induced by performance feedback information. It was expected that acute stress in senior computer interaction would be higher when they receive negative performance feedback than when they receive positive performance feedback.

Chapter 5 examines the effect of complexity and induced expected task performance on acute stress within the Alzheimer domain. The Alzheimer domain was chosen because it is more personally emotionally related to seniors than the art domain. Two similar Alzheimer text retrieval systems were engineered with the exception that one provided low complexity, and the other provided high complexity. The research problem and the design of the experiment were similar to the design of the experiment in chapter 4.

Chapter 6 discusses the findings from the previous chapters. Requirements for NARRATOR are presented, before a more general overview is given. Implications for practice and future research are made, and a conclusion of the thesis is given.

Chapter 2: The design and experimental evaluation of an accessible retrieval system for seniors[2]

2.1 Seniors and information accessibility

It is desirable that public information is accessible for all people; inexperience with current retrieval systems should not be a hindrance to gain access to the information. Seniors have, however, been found to face difficulties learning and using new media (see Charness & Holley, 2004; Docampo Rama, 2001; Freudenthal, 2001; Mead et al., 1999; Mead et al., 2000; Morrell et al., 2002; Xie, 2002), i.e. there are problems with accessibility for seniors. To support successful senior information retrieval, systems should be engineered with reference to the knowledge and capabilities of seniors, and it should be remembered that new media have no functionality gain when a task can be performed more efficiently, effectively and satisfactorily with old media like books.

A thorough understanding of the knowledge and capabilities of the operator, the task environment and task constraints is emphasized by cognitive engineering. Cognitive engineering is an applied cognitive science and was first introduced by Norman (1982), elaborated by Hollnagel and Woods (1983) and Rasmussen (1986). Cognitive engineering is an iterative process until critical accessibility problems in the design are solved. Cognitive engineering aims at compatibility between user and system on at least a number of aspects. First, the system functions and interface elements should be compatible with the users' goals and tasks, and the number of functions should be adequate for the task (i.e. no more than necessary). Second, the system concepts should match the concepts the user employs. Third, because of the user's limited resources, interaction should be minimal, which brings about simplicity. To design a system compatible with a wide range of users with respect to goals and tasks, concepts and limited resources, paper prototyping can be used.

Paper prototyping is a technique that can support system design that enhances effectiveness and efficiency. Building prototypes on paper and testing them with real users has also been called low-fidelity prototyping (Rettig, 1994). Using low-fidelity paper prototypes is a cheap and time saving technique to make many iteration steps in interface design. Because paper prototypes can be redrawn with little cost, they are very effective at the beginning of the development process, because they make it possible to try out many design alternatives (Wickens et al., 2004). Paper prototyping can be effective throughout the product development cycle (Virzi et al., 1996), and leads to almost the same quantity and quality of critical user statements as with computer prototypes (Sefelin et al., 2003).

[2] This chapter is based on Nap, De Greef and Bouwhuis (2005).

2.2.1 Paper prototype study of an art image retrieval system

A paper prototype study was conducted to explore which system functions, interface elements, and system concepts are necessary for an art image retrieval system that is compatible with the goals of both young adults and seniors and the concepts they employ. The art domain was chosen because it is recognizable and pleasurable for many people; hence it is unlikely to elicit negative feelings that could influence interaction performance.

A paper prototype was made of an art image retrieval system, which was used to elicit behaviour from representative users, both young adults and seniors, in tasks intended to provide valid information of the art domain. From the paper prototype study it appeared that the systems' functions should provide the user the possibility to search for images of paintings by the name of the painter, the title of the painting, the time frame of creation, the subject, and style of the painting. The system concepts were matched with the concepts the users employed, e.g. in the first paper prototypes 'Period' was changed into 'Style' to search for images of paintings from a certain art period, style or movement. To maintain simplicity, only the necessary concepts and functions were implemented in the design. No effort was made to beautify the design of the prototype with e.g. colour and fancy buttons. This is because the prototype is developed to provide a *functionality* gain compared to the current situation, and for future improvements of the prototype it is essential to avoid superficial aspects in design that may have a harmful influence on the functionality of the system. The final paper prototype was used to engineer a working computer prototype (see figure 2.1).

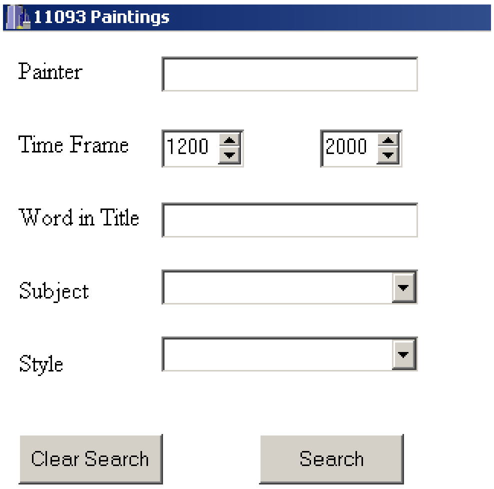

Figure 2.1: Working prototype with which users can search for art images by filling in search words, selecting a time frame, or by selecting items from the drop-down menus. After 'Search' is pressed, the art images appear on the right of the screen (see appendix, figure A.1). The 'Clear Search' button deletes all images and used search words.

2.2.2 Experimental evaluation 1

To gather knowledge about the accessibility[3] of the first prototype for both young adults and seniors, to gather information for possible improvements, and to study whether the prototype provides a functionality gain in comparison with the current situation, an evaluation experiment was performed in which the prototype was compared with two existing art image retrieval web sites. The first was the National Gallery of Art (NGA) web site (NGA, 2004), which uses similar interface elements and concepts as the prototype; however, the interaction is more complex in the sense that there are multiple information layers. The second was the Tigertail Virtual Museum (TVM) web site (TVM, 2004), which uses different interface elements and concepts, features a more complex interaction, and there are multiple information layers. See appendix figures A.2 and A.3 for screen captures of the NGA- and TVM search interface.

The participants did not receive any additional training, beyond the computer experience they already had. Accessible systems should provide the user with adequate affordances, commensurate with their knowledge and experience, to perform simple tasks. It is expected that the task performance of young adults and seniors working with the developed prototype will be more efficient, effective and satisfactory than with the other two systems.

[3] If high accessibility can be provided, the developed prototype could serve as a retrieval platform for health care related information (see NARRATOR, chapter 1.2, this thesis).

2.2.3 Method

Participants

Nine young adults (4 female, 5 male), and seven seniors (7 female), participated at the usability lab of the Technische Universiteit Eindhoven and were paid €15, - (≈ $19,-). The mean age of the young adults was 22.7 (*SD* = 2.2) and of the seniors 63.1 (*SD* = 5.2). For the selection of the participants, use was made of the IPO subject pool of the Technische Universiteit Eindhoven. All young adults were familiar with the English language and had experience with PCs and Internet search engines. The seniors were also familiar with the English language and had little to intermediate experience with PCs (e.g. only e-mail). All young adults were university students, and three of the older adult participants had a higher education than high school (2 college & 1 university). Computer experience was measured by six questions, one open, two with three alternatives and three with four alternatives (example of a four alternative item: "Do you use Internet search systems?", 1. No, never, 2. Yes, less than once a month, 3. Yes, at least once a month, 4. Yes, at least once a week).

Material

The experiment was conducted on a PC with a Pentium 4 processor and with a 15” flat panel screen (1024x768-pixel resolution) and a 2.5 MB/s Internet connection. A standard QWERTY keyboard and optical mouse were used as input and interaction devices. The NGA web site, the TVM web site and the prototype were used as the platforms on which the participants performed the search tasks. The tasks were written on paper. A stopwatch was used to measure task time.

Tasks

Three matching sets of twelve search tasks had to be completed on the three systems. The tasks in each set concerned the name of a painter, the title of a painting, the subject of a painting, the style of a painting, the year of creation and a combination of painter name and subject, and painter name and title. There were six compound- and six single tasks. A single task implied using the search interface once and concerned only one type of information (e.g.: "Find a painting of a dog"). An example of a compound task: "Find a painting of Rembrandt and a painting of Renoir".

Dependent variables *(ISO 9241-11, 1998)*

Accessibility and usability are similar. Increased accessibility promotes an increase in effectiveness, efficiency and satisfaction for people who have a wide variety of capabilities and preferences (see Gulliksen & Harker, 2004). Therefore, the ISO 9241-11 (1998) norm for usability is used as a measure for accessibility.

Effectiveness was measured by the number of search tasks performed correctly. Efficiency was measured by total task time. When a participant was unable to perform a task within the time limit of the specific task, the time limit was coded as the efficiency score. Satisfaction was measured by the IBM satisfaction questionnaire (Lewis, 1995), translated into Dutch. The questionnaire contained nineteen usability items that had to be rated on a scale, running from one to seven (an example item: "I feel comfortable using this system").

Design

To be able to compare the difference in accessibility between systems only one of the following conditions could be varied at a time: user, task, context, and the product (see Gulliksen & Harker, 2004). Because of this, a within-subjects design was used. The presentation order of the three task sets and the three systems was counterbalanced, using a 3 x 3 Latin Square.

Procedure

First, a questionnaire was given to collect the participants' characteristics (such as age, level of education, and computer experience). Also, information was given about the experimental procedure. After this, the participants were given a short introduction about the search tasks they had to perform; this was followed by the twelve search tasks. There were time limits for the search tasks, one minute for the single tasks and two minutes for the compound tasks. This was followed by the satisfaction questionnaire. The participants had a five-minute break to relax and after that, the procedure was repeated for each of the other two systems. An informal debriefing followed with the participants to elicit some background information of computer use in daily life. In the debriefing special care was given to the seniors, in the sense that any possible emotional distress during the experiment was reduced by restorative communication. Information was given about the nature and goal of the research, and the participants were told that their task performance data would be used to improve the retrieval system under study. The total experiment took about one hour and forty-five minutes per participant.

2.2.4 Results

It was planned to employ nine older adult participants, however, the experiment had to be stopped after the seventh participant. A high stress experience was observed, the older adult participants showed high levels of arousal and almost all actions had negative results, which indicated a (stress) state of hopelessness (Ursin & Eriksen, 2004). Three of the older adult participants were not able to perform one task on all three systems within the time limit; eventually they agreed to stop the experiment. This resulted in much missing data that distorted the tasks time means. Therefore, only the number of tasks correct are presented (see table 2.1), which gives an impression of the task performance. In the end, all participants were willing to cooperate with future research. Because most older adult participants stopped beforehand, no satisfaction data were collected.

Table 2.1: Average number of tasks completed correctly (*M*) and their standard deviations (*SD*) of the seniors (*N* = 7) on the three systems

		Prototype	NGA	TVM
Number of Tasks correct	*M*	1.29	1.00	0.86
(max.=12)	*SD*	1.89	1.15	1.46

For the data of the young adults, a repeated-measures ANOVA was performed on total tasks correct, total task times and the total satisfaction ratings for the tasks with system as the within-subjects factor. See table 2.2 for an overview of the young adult task performance.

Table 2.2: Average task performance (*M*) and their standard deviations (*SD*) of the young adults (*N* = 9) on the three systems

		Prototype	NGA	TVM
Number of Tasks correct	*M*	11.33	10.44	9.00
(max.=12)	*SD*	1.00	0.58	0.41
Task Time	*M*	460.78	645.78	817.22
(in seconds, max.=1080)	*SD*	85.38	142.30	96.91
Satisfaction Ratings	*M*	5.62	5.23	3.52
(on a 7 point scale)	*SD*	0.29	0.22	1.27

A significant difference was found between the three systems on total tasks correct [$F(2, 6) = 15.1$, $p = .005$]. Another significant difference was found between the systems on total task times [$F(2, 6) = 169.7$, $p = .000$]. Finally a significant difference was found between the systems on the total satisfaction ratings [$F(2, 6) = 22.1$, $p = .002$].

Pairwise comparisons with a Bonferroni correction (see table 2.3) indicated no significant effect of total tasks correct between the prototype and the NGA web site. Significantly more tasks correct were performed on the prototype than on the TVM web site and significantly more tasks correct were performed on the NGA web site than on the TVM web site.

The participants were significantly faster in performing the search tasks on the prototype than on the NGA web site and the TVM web site. Also, total task times were significantly lower for the NGA web site than for the TVM web site.

On the satisfaction ratings, no significant effect was found between the prototype and the NGA web site, but the participants were more satisfied with the prototype and NGA web site than with the TVM web site.

Table 2.3: Pairwise comparisons of the three conditions on task performance with a Bonferroni correction, significant differences on the 0.05 alpha level are marked with an asterix.

	Prototype > NGA	Prototype > TVM	NGA > TVM
Tasks Correct	$t(8) = 1.315, p = .675$	$t(8) = 3.882, p = .014$*	$t(8) = 3.040, p = .048$*
Task Time (seconds)	$t(8) = 3.844, p = .015$*	$t(8) = 7.411, p = .000$*	$t(8) = 3.717, p = .018$*
Satisfaction Ratings	$t(8) = 1.654, p = .409$	$t(8) = 4.698, p = .005$*	$t(8) = 4.549, p = .006$*

2.2.5 Discussion

Young experienced users clearly benefit from the developed prototype in retrieving images according to the instructions; the seniors with less experience did not. Most seniors could simply not operate the existing exemplars. But they did also not benefit from the developed prototype in retrieving images according to the instructions. Some seniors did not carry on with the tasks and showed high levels of stress. Similar results were found by Birdi and Zapf (1997), who found that older workers show a more negative emotional reaction to errors in computer-based work than younger workers; they were less likely to try to solve a problem entirely in response to an error situation.

The results illustrate the importance of involving seniors in the design process from an early stage (see also Hawthorn, 2006), given that younger adults were very capable of operating the first working prototype. This has practical implications for designers: ignoring seniors in the design process early, and experimentally evaluating only the final system with seniors, would be costly in terms of time and money for redesign and improvement of the system. The working prototype under study, however, can still be easily amended, experimentally evaluated and improved.

In view of the very low performance of the seniors, the user interface design studied in this first evaluation must have contained flaws. To better understand the problems they were facing and to gather design recommendations for new prototypes and to evaluate the new prototypes, we analysed the errors made, amended the user interface design, and performed a second evaluation.

2.3.1 Error analysis & redesign

It appeared not difficult to classify the errors by the screen location in which they occurred. A rough error analysis, considering all the erroneous search words that were used by the seniors to perform the tasks, shows that 59% of the errors were made at the 'time frame' and 'drop-down menu's'. Of the errors, 27.3% could have been caused by taking the tasks to the letter, e.g. the task "Find a portrait of Doré (Gustave)" resulted in the search string "Doré (Gustave)". The parentheses symbol was not recognized by the system and consequently no results were shown. This classification provides insufficient information about possible causes and remedies.

Many different approaches for human error classification have been developed (Embrey, 1986; Hollnagel, 1998; Reason, 1987). We used a classification scheme, comparable to SHERPA (Embrey, 1986) and influenced by human dialogue models (Bouwhuis, 1991), which provided us with an insightful overview of the poorly performed tasks. Three different aspects of an erroneous search word were considered for classification[4]. First: is there a match between the task and the search word used? If not, then it is classified as a task comprehension error. For example, the user selected still life paintings in the drop-down menu and started searching, while the search task concerned a painting of people. Second: if there is a match, is it located on the correct object? If not, then it is classified as a location error. For example, the user typed the search word ‘still life’ in the painter textbox and started searching. Third: if there is a match and located on the correct object, is it recognized by the system? If not, then it is classified as a system comprehension error. For example, the user miss spelled ‘Rembrandt’, typed ‘Rembant’ in the painter textbox, and started searching. Combinations of the three aspects are possible.

About 63% of the errors made by the seniors were classifiable in this study (110 errors of 167 used search words). The errors were quite diverse compared to the young adults for which more than 97% of the errors were classifiable (42 errors of 211 used search words). An overview is presented in table 2.4.

[4] It should be noted that the errors occured because of interface design flaws. The error classification is not intended to blame the user.

Table 2.4: Classifications of most frequent errors occurred for the seniors and young adults

Percentage of Erroneous Tasks (%)		***Error Type***	***Description***	***Consequence***	***Illustrative Remedies/Redesign***
Seniors	**Young adults**				
19.5	30.9	Task Comprehension Error	Type or select wrong information for task	Wrong results shown	Offer Help-file with examples of search tasks
18.9	35.7	System Comprehension Error	Type wrong information for system	No results shown	Restricted interaction, only selection possible in drop-down menu, add Dutch to English database
10.0	2.3	System Comprehension & Location Error	Type wrong information for system on wrong object	No results shown	Restricted interaction, only selection possible in drop-down, ignore non-database terms
8.3	0.0	Location Error	Type right information for task on wrong object	No results shown	One search box, restricted interaction
6.5	21.4	Location & System Comprehension Error	Type partly correct information for task on right object	No results shown	Ignore non-database terms
0.0	7.1	Location & System Comprehension Error	Type partly correct information for task and system on right object	Wrong results shown	Offer Help-File with examples of search tasks, ignore non-database terms
63.2%	97.4%				

2.3.2 Three redesigns

As a result of the error classifications and the illustrative remedies, three different redesigns were constructed, intended to overcome task comprehension-, location- and system comprehension errors. The first one was a closed interface (see figure 2.2), in which interaction is restricted in four steps to support the user during interaction and dialogue steps. First, the user has to select a category, and then fill in a search word. After that, the button 'Search' can be pressed and after the search, 'Clear Search' can be used. The second was a help interface (see figure 2.3) offering knowledge in the form of an assist function that could compensate knowledge deficiencies (see De Greef & Neerincx, 1995; Neerincx & De Greef, 1998), and interaction is partly restricted. Finally there was a one-box interface (see figure 2.4), in which users cannot type any information for a task and the system in a wrong box.

In all three redesigns the font size was 14 (in points), compared to 12 in the first prototype and the height of the ‘Search’ and ‘Clear Search’ buttons was enlarged by 23%, which resulted in buttons with a surface area as large as 4000 pixels. The buttons that can be used to browse through the art images after ‘Search’ is pressed, which had a 25 pixel height and 46 pixel width, were replaced by buttons with a 65 pixel height and 118 pixel width. The arrows on these buttons were replaced with the text ‘Volgende’ (next) and ‘Vorige’ (previous) with an 18-point font size. See appendix figure A.4 for a screen capture of the help interface after ‘Search’ is pressed. Furthermore, all text in the three redesigns was in Dutch. These adjustments were made because some of the classified errors could have been caused by visual perception problems and language misunderstandings, in fact, these adjustments correspond to requirements originated by other applied cognitive ageing researchers. A recent Cognitive Engineering method (CE+) proposed by Neerincx and Lindenberg (in press) suggests that relevant guidelines or requirements must be addressed from the beginning of the design of an interface. See Czaja and Lee (2003), Fisk, Rogers, Charness, Czaja and Sharit (2004), Making Your Web Site Senior Friendly (2006), Mead et al. (2000), Morrell et al. (2002), and Welford (1985), for in-depth overviews of computer system design requirements for older adult users, based on theoretically driven research.

Figure 2.2: Closed interface (translated from Dutch to English: ‘Druk op Zoek als u klaar bent met invullen’ means ‘Press Search when ready filling in’, ‘Schilder’ means ‘Painter’, ‘Jaartal’ means ‘Year’, ‘Titelwoord’ means ‘Word in Title’, ‘Onderwerp’ means ‘Subject’, ‘Stijl’ means ‘Style’, and ‘Zoek’ means ‘Search’)

Figure 2.3: Help interface (for translation see Fig. 2.2, ‘Hulp’ means ‘Help’, ‘Wissen’ means ‘Delete’ and on the bottom of the screen a precise description is given for how to perform a search task)

Figure 2.4: One-box interface (for translation see Fig. 2.2 & 2.3)

2.4 Experimental evaluation 2

A second evaluation experiment was performed in which the three redesigns were compared. It is expected that the task performance of seniors would be better than with the initial system design, and the user experience would be less stressful.

2.4.1 Method

Participants

Nine seniors participated at the usability lab of the Technische Universiteit Eindhoven and were paid €15,- (≈ $19,-). The participants differed from the first study and were all women (mean age = 65.78, *SD* = 6.96), with little to intermediate experience with PCs. Seven of the participants had a higher education than high school (5 college & 2 university). For the selection of the participants, use was made of the IPO subject pool of the Technische Universiteit Eindhoven.

The method of the study was the same as for the first evaluation, except for changes in the systems and task formulations. The three redesigns were compared and redundant information in the task formulation was removed, e.g. the addition of a surname to find a painting of a particular painter.

2.4.2 Results

All participants were able to perform search tasks on the three systems, most often within the time limit, and no stress was observed. See table 2.5 for an overview of the task performance.

Table 2.5: Average task performance (*M*) and their standard deviations (*SD*) (*N* = 9) on the three redesigns

		Closed	**One-Box**	**Help**
Number of Tasks correct	*M*	8.78	6.56	8.22
(max.=12)	*SD*	1.92	3.71	3.19
Task Time	*M*	708.89	751.33	706.22
(in seconds, max.=1080)	*SD*	127.35	263.24	222.30
Satisfaction Ratings	*M*	6.22	4.19	5.86
(on a 7 point scale)	*SD*	0.83	2.04	1.23

A repeated-measures ANOVA was performed on total tasks correct, total task times and the satisfaction ratings for the tasks with system as the within-subjects factor. No significant effects were found on total tasks correct and total task times between the systems. A significant main effect was only found on the satisfaction ratings [$F(2, 6) = 7.29$, $p = .006$]. Pairwise comparisons with a Bonferroni correction[5] indicated that the participants were significantly more satisfied with the help interface than with the one-box interface [$t(8) = 3.11$, $p = .043$].

2.4.3 Discussion

The performance of the seniors with the redesigns was much better compared to the initial system design. It is not clear what mechanisms are responsible for the enhanced performance with the redesigns in comparison with the first prototype, although it does show that cognitive engineering of interfaces is successful. In the general interface guidelines for older adults (Czaja & Lee, 2003), nineteen guidelines are presented and two of the guidelines state that small targets and characters (fonts < 12) should be avoided and that the size of icons should be maximized. The common system design characteristics of the three interfaces are large fonts and buttons, and the use of the native language; we expect that these characteristics have an impact on the performance increase. According to this research, the surface area of an icon might need to be as large as 4000 pixels (e.g. 40 height x 100 width). In addition, fonts smaller than 14 might need to be avoided, which is consistent with Echt's (2002) interface design guidelines for seniors.

From the participants' comments and expressive behaviour, it appeared that they were satisfied to work with the three systems, in particular the closed interface. The help interface was significantly more satisfactory than the one-box interface, but overall, the differences in task performance between the three interfaces were small. According to Czaja & Lee (2007), older adults are more likely to make errors when interacting with technical systems and rely on help systems more for error recovery. It is interesting to note that the help function was hardly used, so the system's functions and design without help are very likely to be sufficient for effective, efficient, and satisfactory information retrieval. On the other hand, if the seniors had more time to perform a search task (e.g. 5 minutes) they would have had the opportunity to gradually read the available help information. Overall, the systems gave the seniors a clear sense of mastery and control. After the experiment, they started searching art images they personally favoured and three of the participants indicated that they would like to have the closed- or help interface at home. The participants suggested that the system could be improved by the possibility of enlarging the art images, and by an image export function to a text editor or e-mail program, and by additional information about in which museum the painting is located, so you could go and visit to see the real painting. Possible limitations of the experimental evaluation are the small sample size, the lack of comparison with the first prototype and the carry-over effect that is always present in within-subjects repeated-measures research.

[5] With a LSD correction, the participants were significantly more satisfied with the help interface and closed interface than with the one-box interface.

With the redesigns, the seniors performed as well as the young adults with the TVM web site, and they were more satisfied with the closed interface than the young adults with any of the three systems were. Although a large performance increase was found for the seniors, seniors were still slower in performing the search tasks on the redesigns than were the young adults on the first prototype. The task time performance scores of the seniors are much more widely spread than the scores of the young adults, which indicate that there are large individual differences in efficiency. This is in line with other studies (see Czaja & Lee, 2007), which indicated considerable variability in performance among older people (60-75 years). Overall, the seniors that participated in the second evaluation experiment had a higher education than the seniors in the first experimental evaluation. Nevertheless, it is unlikely that the level of education contributed to the difference in task performance between both experimental evaluations, because the seniors in both experiments had similar experience with PCs, and the seniors in the first experiment that did had a higher education than high school were also not able to perform search tasks on the first prototype.

2.5 Experimental evaluation 3

The interface design of the previous experimental evaluation is usable for seniors. A third evaluation study was performed to study if the interface provides accessibility for all, and to overcome limitations of the previous evaluation experiments, e.g. the small sample size and the carry-over effect. With a wide range of users, the first developed prototype was compared with the help system (redesign), without the help function. The help function was removed, because most seniors in the previous experiment did not use it, and because of limited available time to read through the descriptions of how to perform a search task. It is expected that task performance on the first developed prototype will be lower than on the redesigned system.

2.5.1 Method

Participants

Eighty-one visitors of the 'Technology Management Public Day' voluntarily participated at the audio visual lab of the Technische Universiteit Eindhoven. All participants were quasi-randomly assigned, allocated by available computer, to one of two groups. Thirty-one participants were deleted from the sample. These included participants who did not finish the experiment, and people that participated more than once. If the age and gender of a successive participant was the same as the previous participant and the time between participation less than two minutes, the second or third dataset was disregarded. From the remaining fifty participants, twenty-five (17 male, 8 female) participated in the first prototype condition. The mean age was 20 (*SD* = 15.7), min. age was 7 and max. age was 62 (median = 12).

Twenty-five (17 male, 8 female) participated in the redesign condition. The mean age was 15.4 (*SD* = 12.4), min. age was 7 and max. age was 62 (median = 12).

Material

The experiment was conducted on a PC with a Pentium 4 processor and with a 15" flat panel screen (1024x768-pixel resolution). A standard QWERTY keyboard and optical mouse were used as input and interaction devices. The first developed prototype (see figure 2.1) and redesign (see figure 2.3), without the help function, were used as the platforms on which the participants performed the search tasks. The tasks were presented on the flat panel screen, task time and task correct was automatically logged by the program.

Tasks

Five tasks had to be completed. The tasks in both groups were matched and concerned the name of a painter, the title of a painting, the subject of a painting, the style of a painting, and the year of creation.

***Dependent variables** (ISO 9421-11 norms for usability)*

Effectiveness was measured by the number of search tasks performed correctly, and efficiency was measured with total task time. When a participant was unable to perform a task within the time limit of the specific task, the time limit was coded as the efficiency score.

Design

A between-subjects design was used.

Procedure

First, a questionnaire was given to collect the participants' age and gender. After this, the participants were given a short introduction about the search tasks they had to perform, and this was followed by the five search tasks. There was a one-minute time limit each task. After the tasks, the participants were thanked for their participation. The total experiment took about five minutes per participant.

2.5.2 Results

A one-way ANOVA was performed on total tasks correct and total task times with system as the between-subjects factor (see table 2.6 for averages and standard deviations).

No significant difference was found between the systems on total tasks correct [$F(1, 48) = 2.44$, $p = .125$], and no significant difference was found between the systems on total task times [$F(1, 48) = 2.16$, $p = .148$].

Table 2.6: Average task performance (M) and standard deviations (SD) on the first prototype and redesign

		First Prototype (n = 25)	Redesign (n = 25)
Number of Tasks correct	*M*	3.36	3.96
(max.=5)	*SD*	1.52	1.17
Task Time	*M*	154.20	122.47
(in seconds, max.=300)	*SD*	90.70	58.59

Because each task represented a functionality of the system, the differences between the task performance for the different tasks were compared. A one-way ANOVA was performed on the total task times and total tasks correct of the five tasks performed. The total task times of the Year question on the redesign was significantly lower than on the first prototype, [$F(1, 48) = 11.3$, $p = .002$], with an effect size[6] (partial η^2) of .190 and an observed power of .908. The number of tasks completed correctly of the Year question on the redesign was significantly higher than on the first prototype, [$F(1, 48) = 12.4$, $p = .001$], with an effect size (partial η^2) of .205 and an observed power of .932. The number of tasks completed correctly of the Style question was significantly higher on the redesign than on the first prototype, [F(1, 48] = 4.75, $p = .034$], with an effect size (partial η^2) of .090 and an observed power of .570.

2.5.3 Discussion

The analyses show that overall the first prototype and the redesign do not differ in effectiveness and efficiency for a wide range of users, but they do differ when the task performance of some of the different tasks are compared. The functionality to search a painting by year of creation and style is flawed in the first prototype, this is consistent with the error analysis that was performed after experiment 1, where it was found that most errors were made at the 'time frame' and 'drop-down menus'. As mentioned, to compare the difference in accessibility between systems one must only vary one of the following terms at a time: user, task, context, and the product. In this research, the user and the product were varied together, which makes it difficult to draw conclusions about the accessibility of the redesign. A possible limitation of the study is the number- and difficulty of the tasks that had to be performed.

[6] An η^2 of .01 indicates a small effect, .06 a moderate effect, and .14 a large effect (Cohen, 1988).

2.6 General discussion and conclusion

The first experiment shows that even a modest amount of cognitive engineering can bring about a significant performance enhancement for the young adults. The young adults were more efficient in performing the search tasks on the cognitively engineered prototype than on the NGA- and TVM web sites. This result is in line with the results of Gerhardt-Powels (1996); she found that a cognitively engineered interface is superior in measures of reaction time, accuracy, workload and preference, compared to two other interfaces that were not designed with explicit reference to cognitive engineering principles. The experiment shows that seniors with some computer experience may still face considerable difficulties in performing tasks with the tools available on the web.

A more elaborate form of cognitive engineering includes a careful error analysis, improvement of the design and perhaps adding special functions to support users' cognitive functioning. The results from the second experiment show that some minor adjustments can have a tremendous effect on performance of seniors. Other strong benefits of this type of cognitive engineering of interfaces have been found in statistical analysis (De Greef & Neerincx, 1995), and railway traffic control (Neerincx & De Greef, 1998). The second experiment shows that seniors are capable of operating an image retrieval system when it is cognitively engineered according to the methods described in this paper.

As mentioned by Hawthorn (2000), there are almost no studies on what makes an interface usable for older adults. In this research, both young adults and seniors performed better on a category based image retrieval system than on a one-box image retrieval system. Furthermore, seniors seem to benefit from the use of their native language in the user interface and system, large fonts and large buttons, and the use of concise task formulations in the evaluation experiment.

A careful and systematic iterative design of a minimal interaction system can produce a vast step towards high accessibility and a user experience with low or no stress. This design strategy requires the early involvement of seniors in the design process. The final system used in this study is more readily usable for seniors and young adults, compared with existing exemplars. For seniors to perform search tasks on the redesign, no additional training is needed, a good level of performance is shown with standard interaction devices, and there is little or no stress in using the application. Although this is a valuable improvement, other related factors such as stress should be investigated more thoroughly and systematically, in search for more satisfying usage in the long run. Stress appears to be an important factor in senior computer interaction, but little is known about the factors that influence stress in senior computer interaction (see Hamborg & Grief, 2003; Hawthorn, 2000).

Acknowledgement

Thriloke Thakur performed the paper prototyping studies, and the name and functionality principles of a closed interface were originated by him, in 2003.

Chapter 3: The effect of complexity on acute stress in senior computer interaction

3.1 System complexity and acute stress

The 'beating up the printer' scene in the movie "Office Space" (Judge, 1999), illustrates how technology can drive people mad. Although the scene might seem a little exaggerated, annoyance and stress is easily observable in people who work with computers. If experienced computer users can be stressed and annoyed because the computer output does not match their expectations (discrepancy between a set and a real value), what about inexperienced seniors who have been raised in a different technological generation (Docampo Rama, 2001)?

The western population is aging, and these seniors are faced with rapidly changing technologies in their work and home environment. Because of this, it is important to study the possible effects of new media on stress in seniors, and understand what underlying mechanisms contribute to stress. In the previous chapter (experiment 1, this thesis), it was shown that stress might indeed be an important factor in senior computer interaction. Acute stress was observed when seniors performed tasks on a suboptimal image retrieval system, while after redesign no stress was observed. In spite of these findings, it is far from clear exactly why the seniors were stressed and how stress can be reduced in senior computer interaction.

According to Ursin and Eriksen (2004), the stress response is a general, unspecific alarm response occurring whenever there is a discrepancy between a set value (what is expected or the normal value), and an actual value (what is happening in reality). The stress response is dependent on the available responses or resources (Ursin & Eriksen, 2004). In the interaction between humans and technology, available resources are present in people, in the interaction with the interface, and in the systems' functional technology, independent of the interaction. It is likely that seniors are stressed during computer interaction when the complexity of a system is high. If the subjective probability of eliminating the discrepancy between a set- and an actual value is low because of high complexity, a subjective stress experience and physiological stress response is likely to happen. A similar relation has been found between subjective stress and absolute task difficulty; subjective stress increases when absolute task difficulty increases (Callister, Suwarno & Seals, 1992). To reduce stress, support can be offered by accessible systems. With an accessible system, the user has the tools (i.e. control) to cope with task complexity demands. Control and coping are not identical phenomena. According to Ursin and Eriksen (2004), control is an acquired perceived high probability of a given response outcome, regardless of the value of the outcome, whereas coping is the acquired expectancy that most or all responses lead to a positive result and consequently reduce arousal.

The experiment presented here was conducted to examine the extent to which complexity causes stress in senior computer interaction. In the previous chapter, it was shown that seniors show high levels of acute stress while performing search tasks on a high complexity system. Furthermore, it was shown that no stress is observed when seniors perform tasks on a low complexity system. Shortcomings of the previous experiment were that stress was observed in seniors, but not objectively measured, and no direct comparison was made between the accessibility of the low- (redesign) and high complexity systems (first prototype) for seniors.

3.2 Experiment

3.2.1 Method

Participants

Twelve seniors were invited for participation in the usability lab of the Technische Universiteit Eindhoven. For the selection of the participants, use was made of the IPO subject pool of the Technische Universiteit Eindhoven. In total, nineteen seniors signed up for participation from which three cancelled participation, two forgot the information letter and were not able to find the usability lab on the campus. Two seniors used blood pressure (BP) medication, were withdrawn from the sample, and were paid €15,- (incl. travel expenses). Of the remaining twelve seniors, there were six males and six females (mean age = 66.92, *SD* = 5.52), and they were paid €15,-. Criteria for participation were: familiarity with the English language, subjectively healthy, and having intermediate experience with PCs (beside Microsoft Word, they also used other programs). They were asked if they had diabetes, and if they used BP medication, heart medication, or both. None of the remaining participants had diabetes, used BP medication, heart medication, or both.

Material

The experiment was conducted on a PC with a Pentium 4 processor and with a 15” flat panel screen (1024x768-pixel resolution). A standard QWERTY keyboard and optical mouse were used as input and interaction devices. The low- and high complexity systems were used as the platforms in which the participants performed the search tasks. The low complexity system was identical to the help interface, without the help button (see chapter 2, figure 2.3). The high complexity system was identical to the first prototype (see chapter 2, figure 2.1). The main differences between the low complexity and high complexity systems were the font and button sizes, and the text that was presented on the buttons. The font size on the buttons in the low complexity system was 14 (in points) compared to 12 (in points) in the high complexity system. The buttons in the low complexity system were 23% larger than in the high complexity system. In addition, the text on the buttons was in Dutch in the low complexity system compared to English in the high complexity system. The tasks were presented on the screen (see figure 3.1), and effectiveness and efficiency was automatically logged by the program. BP and heart rate (HR) was measured with an arm cuff connected to an automatic BP measurement device (Omron M6), using the oscillometric method, and is clinically validated according to the International Protocol of the European Society of Hypertension (Topouchian, El Assaad, Mohamed, Orobinskaia, Ludmila, El Feghali, Ramzi, Asmar & Roland, 2006).

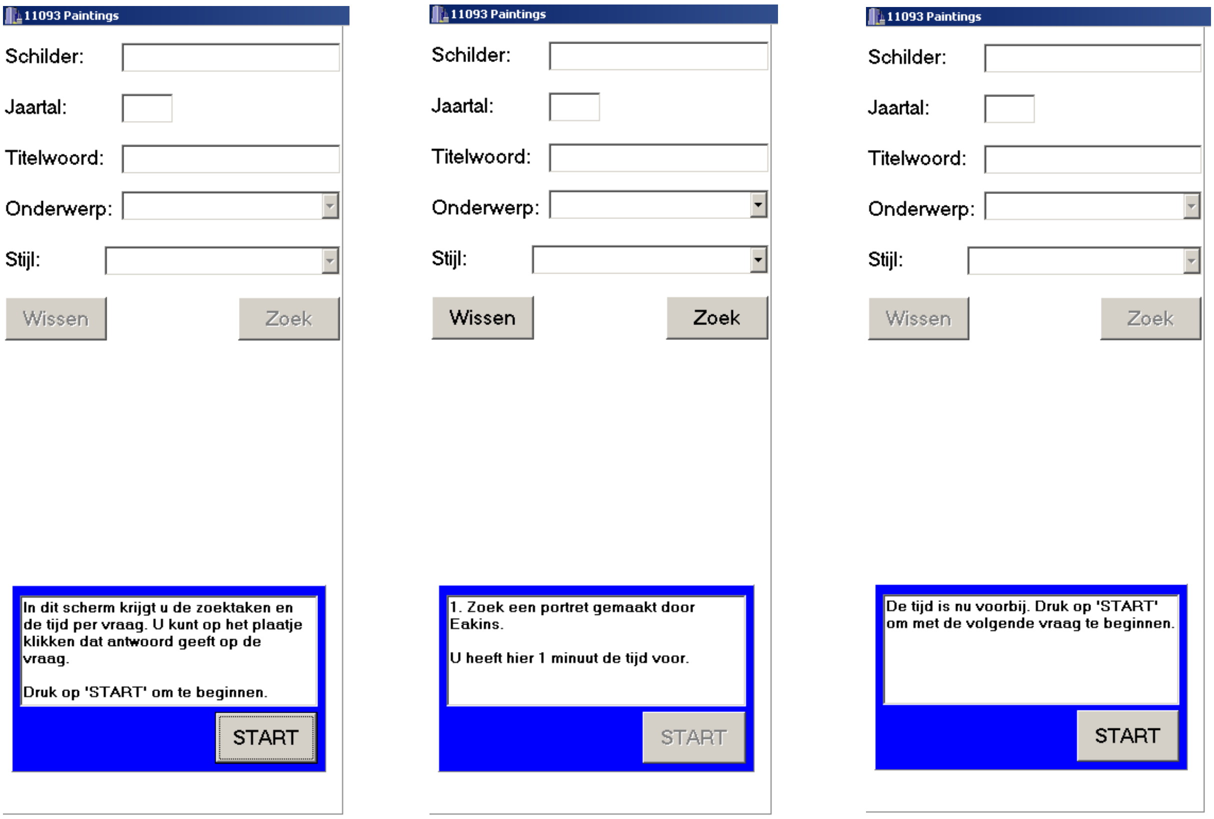

Figure 3.1: Task presentation during experiment on low complexity system, bottom left (inside blue field): introduction about how to perform a task, bottom middle: presentation of a task, bottom right: end of a task and instruction to proceed with the next question

Tasks

Two matching sets of ten search tasks were to be completed on the two systems. The search tasks in each set concerned the name of a painter, the title of a painting, the subject of a painting, the style of a painting, and the year of creation. An example of a search task was: "Find the painting 'Peppermint'". After a participant filled in or selected a search word on the image retrieval system (e.g. 'Peppermint' at the 'Word in title' fill-in box (translated into Dutch: 'Titelwoord', see figure 3.1), the search button (Zoek) could be pressed. Then matching art images were presented to the right of the search interface. Participants could browse through the images. When they thought they knew the correct answer to a search task the corresponding art image was to be clicked on.

Dependent variables

Acute Stress

Subjective stress was measured by having participants fill in a rating scale on ten attributes, five positive and five negative. The positive attributes: concentrated (geconcentreerd), motivated (gemotiveerd), happy (gelukkig), unconcerned (zorgeloos), and self-confident (zelfverzekerd). The negative attributes: stressed (gestrest), tensed (gespannen), irritated (geprikkeld), nervous (zenuwachtig), and exhausted (uitgeput). These attributes had to be rated on an 11-point rating scale (0 = not at all, 10 = totally). Five of the above attributes, two of the positive (concentrated and happy) and three of the negative (stressed, tensed, and exhausted), were used by Rissén, Melin, Sandsjö, Dohns, and Lundberg (2000). The other five attributes are semantically related to the attributes, and are either synonyms or antonyms of the attributes used by Rissén et al. (2000). The rating technique was developed by Lundberg and Frankenhaeuser (1980), and had been found to correlate with physiological responses, and is still used today (Hjortskov et al., 2004).

Physiological stress (stress response) was measured by systolic- and diastolic BP, and HR, following the ESH recommendations for BP measurement (O'Brien, Asmar, Beilin, Imai, Mallion, Mancia, Mengden, Myers, Padfield, Palatini, Parati, Pickering, Redon, Staessen, Stergiou & Verdecchia, 2003). The mean value of three readings was used for analysis. Systolic BP refers to the pressure of blood in the artery when the heart contracts and diastolic BP refers to the pressure of blood in the artery when the heart relaxes between beats.

Task Performance (ISO 9241-11, 1998)

Effectiveness was measured by the number of search tasks performed correctly. Efficiency was measured with total task time. When a participant was unable to perform a task within the time limit of the specific task, the time limit was coded as the efficiency score. Satisfaction was measured by the IBM satisfaction questionnaire (Lewis, 1995), translated into Dutch. The questionnaire contained nineteen usability items that had to be rated on a scale, running from one to seven (an example item: "I feel comfortable using this system").

Independent variables

Complexity

Low, which was induced by an art image retrieval system (see figure 3.1) that has been validated as a high accessible system for seniors (chapter 2, this thesis), and high (see figure 2.1) that has been validated as a low accessible system for seniors (chapter 2, this thesis).

Design

A within-subjects design was used. The participants performed two matched sets of similar tasks on the two systems. The presentation order of the two task sets and the two systems was counterbalanced.

Procedure

The participants received an informed consent form by letter. On arrival, the signed informed consent was collected and information was given about the experimental procedure. It was explicitly stated that the participants were allowed to stop the experiment whenever they wanted. They had five minutes to relax and to drink some tea. Then BP was measured three times in a row, and subjective stress was measured (baseline measurement). After this, a questionnaire had to be filled in to collect the participants' characteristics. The participants were given a short introduction about the search tasks they had to perform, and had three minutes to relax. BP was measured three times, and subjective stress was measured (pre-measurement). Next, participants carried out the ten search tasks. The tasks were presented on the screen (see figure 3.1). After the participants pressed start, the search interface became active. After the task time limit, the search interface became inactive. The time limit per search task was one minute. This was followed by measuring BP three times, and subjective stress (post-measurement). After this, the satisfaction questionnaire was to be filled in. The participants had a three-minute break, and then the procedure was repeated for the other system. In the debriefing, any possible emotional distress during the experiment was reduced by restorative communication. Information was given about the nature and goal of the research. The total experiment took about one hour per participant.

3.2.2 Results

Stress scale reliability

Cronbach's alpha is a coefficient of reliability, and if the inter-item correlations are high (>0.70), then there is evidence that the items are measuring the same construct. The mean Cronbach's alpha of the five positive subjective stress measurement items was 0.46 (range 0.35-0.53). The mean Cronbach's alpha of the five negative subjective stress measurement items was 0.77 (range 0.54-0.94).

The five positive subjective stress items are disregarded in the following analyses due to a lower Cronbach's alpha than 0.7. Many participants complained about the negative subjective stress item 'exhausted,' and found it unrelated to the task at hand. When the item 'exhausted' was left out in the negative subjective stress items, the mean Cronbach's alpha became 0.81 (range 0.67-0.93), the mean values of the remaining four negative subjective stress items were used as measures of the subjective stress variable.

The effect of complexity on acute subjective- and physiological stress

A repeated-measures ANOVA was performed with complexity (2 levels: high and low) and time (3 levels: baseline, pre and post) as within-subject factors. Within these six conditions, systolic BP, diastolic BP, HR, and subjective stress were measured (see table 3.1 for descriptives).

Table 3.1: Averages and standard deviations of physiological- and subjective stress before (pre) and after (post) the search tasks were performed on the low- and high complexity systems

			Complexity			
			Low		**High**	
		Baseline	Pre	Post	Pre	Post
Sys BP	*M*	135.39	136.39	130.67	135.31	130.36
(mmHg)	*SD*	16.06	19.51	16.78	16.02	13.60
Dia BP	*M*	72.56	74.81	71.33	73.42	71.78
(mmHg)	*SD*	9.12	11.01	10.35	8.79	9.42
HR	*M*	75.53	72.64	70.56	72.89	71.03
(BPM)	*SD*	12.93	10.16	10.19	12.62	11.69
Sub Stress	*M*	1.88	1.67	1.82	1.69	1.96
(max. = 10)	*SD*	1.18	1.27	1.28	1.51	1.25

Complexity did not significantly influence systolic BP, [$F(1, 11) = .210$, $p = .656$], with an effect size (partial η^2) of .019 and an observed power of .070. In addition, complexity did not significantly influence diastolic BP, [$F(1, 11) = .152$, $p = .704$], with an effect size (partial η^2) of .014 and an observed power of .065. Complexity did not significantly influence HR, [$F(1, 11) = .156$, $p = .701$], with an effect size (partial η^2) of .014 and an observed power of .065. No significant influence of complexity on subjective stress was found either [$F(1, 11) = .344$, $p = .570$], with an effect size (partial η^2) of .030 and an observed power of .084.

Time significantly influenced systolic BP, [$F(1, 11) = 5.31$, $p = .042$], with an effect size (partial η^2) of .326 and an observed power of .556., and HR, [$F(1, 11) = 25.1$, $p < .001$], with an effect size (partial η^2) of .695 and an observed power of .995. Pairwise comparisons with a Sidak correction indicated that the participants' systolic BP was significantly lower after performing the search tasks (post) than before performing the search tasks (pre), [$t(11) = 3.279$, $p = .022$]. Their HR was significantly higher at the baseline than before performing the search tasks (pre), [$t(11) = 3.589$, $p = .010$], and their HR was significantly lower after performing the search tasks (post) than before performing the search tasks (pre), [$t(11) = 4.259$, $p = .004$].

Task Performance

A repeated-measures ANOVA was performed on total tasks correct, total task times and the satisfaction ratings for the tasks with complexity (low & high) as the within-subjects factor (see table 3.2 for descriptives). No significant effects were found on total tasks correct, total task times and the satisfaction ratings between the complexity systems.

Table 3.2: Average usability scores and their standard deviations ($N = 12$) on the low- and high complexity systems

		Complexity	
		Low	**High**
Number of Tasks correct	*M*	5.00	3.33
(max.=10)	*SD*	3.67	3.47
Task Time	*M*	496.59	555.71
(in seconds, max.=600)	*SD*	140.37	157.98
Satisfaction Ratings	*M*	4.54	4.05
(on a 7 point scale)	*SD*	1.44	1.75

Systolic BP of participant 7

One of the participants (participant 7) showed an exceptional rise in systolic BP (see figure 3.2). The lowest systolic BP measured (pre1 low) was 165 mmHg, and the highest systolic BP measured (post2 high) was 193 mmHg. The difference between the pre- and post-measurements for the low- or high complexity systems were similar, just like for all other participants.

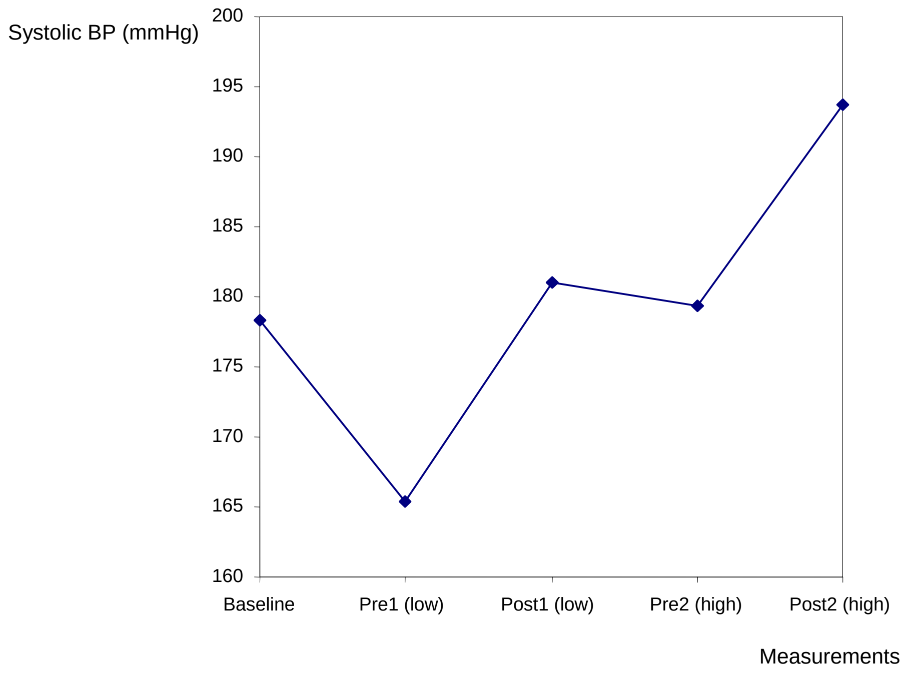

Figure 3.2: Systolic BP of participant nr. 7, before (pre) and after (post) interaction with low complexity (low) and high complexity (high) system

3.2.3 Qualitative results

No stress was observed during the experiment. Not all participants seemed to understand the experimental procedure, in particular the presentation of the tasks on the screen. From the participants' comments and expressive behaviour it appeared that they perceived the task presentation as part of the interface, as if the retrieval system was solely a system to test art knowledge. In addition, the participants thought that they needed knowledge of the art domain to perform the tasks. One of the participants tried to press the buttons on the screen, instead of using the mouse and keyboard. Most participants browsed through the images, instead of refining their search words.

3.3 Discussion

When seniors interact with a high complexity system, their acute subjective- and physiological stress is not higher than when they interact with a low complexity system. Systolic BP and HR were significantly lower after they performed the search tasks, independent of the complexity of the system. This effect can be attributed to the participants sitting (i.e. humans' physiological resting posture) in a chair on the ground for one hour.

No significant task performance difference was found between the complexity of the systems. Consequently, the accessibility of the high complexity system and the low complexity system are similar. Because of this, it is understandable that no differences were found in the seniors' level of subjective- and physiological stress across system complexity. It should be mentioned that all participants had intermediate experience with PCs, whereas the seniors that participated in the experiments in chapter 2 (this thesis) had little to intermediate experience with PCs. Their task performance on the high complexity system was close to zero, whereas the seniors that participated in this experiment performed about a third of the search tasks correctly on the high complexity system.

It is likely that the experiment was flawed by the way the tasks were presented on the screen. Most seniors were confused about the procedure, perceived the task presentation as part of the interface, and browsed through the images instead of using the category-based search options. This could have been caused by age-related changes in attention, which have made it difficult for seniors to switch their attention between competing displays of information or process multiple information formats simultaneously (see Czaja & Lee, 2007). Most seniors have a lot of experience with tests and task presentations on paper, but little experience with automated experimental tests. Automated task presentations should be avoided in senior computer interaction research.

During the experiment, one of the seniors' systolic BP rose considerably. A daytime average BP of 135 mmHg (Sys) and 85 mmHg (Dia) or above is considered to be hypertension (see Simon, 2000), however, most seniors would fall into this category. The low complexity system that was provided to cope with task complexity demands might have been insufficient for this particular senior. The whole experiment could have been physiologically stressful, although he was not subjectively stressed and did not want to stop with the experiment. Anxiety can increase BP by as much as 30 mmHg (Sys), and if this anxiety is an alerting reaction to a doctor, it is commonly referred to as the 'white-coat effect' (O'Brien et al., 2003). In the present study, it may well have been a case of 'white coat effect' (see McAlister & Straus, 2001; O'Brien et al., 2003; Simon, 2000), in which patients have high levels of BP in a clinical setting but nowhere else. Almost 20% of patients diagnosed as hypertensive based on readings in the clinic have entirely normal blood pressures outside the clinic (McAlister & Straus, 2001). Patients are unaware of the 'white coat effect', which could explain why participant nr. 7 was not subjectively stressed. Because his systolic BP was 20 mmHg higher at the baseline measurement than at the first pre-measurement, it likely was a case of 'white coat effect.' If his systolic BP was consistently high (e.g. above 180 mmHg), the experiment could have been risky for his health, and should have been stopped. Researchers who study physiological responses, in particular BP, to certain stimuli, should be aware of the possibility of the 'white coat effect' when conclusions are drawn from the findings. It is likely that high BP values will be measured at the beginning of an experimental condition, because of the 'white coat effect,' and lower BP values will be measured at the end of an experimental condition, when subjects had time to restore allostasis. To measure stimuli specific (direct) BP responses, a continuous BP measurement device, such as 'The Portapres' of TNO could be used, however the technique is subject to various inaccuracies (see O'Brien, Waeber, Parati, Staessen & Meyers, 2001), and the latest model is awaiting formal validation (dableducational, 2007).

In conclusion, no evidence was found that complexity has an effect on acute stress in senior computer interaction. To understand stress in senior computer interaction, other factors, such as 'expected task performance' or 'outcome expectancies' (see Ursin & Eriksen, 2004) of the task at hand, should be considered.

Acknowledgements

Thanks to H. Nap, M.D. – General Practitioner, and R. Lionarons, M.D. – Cardiology, for their advice considering BP measurement methods and interpretation.

Chapter 4: The effect of complexity and expected task performance on acute stress

4.1 Theoretical model of acute stress in senior computer interaction

Today, many people are stressed due to demands of work, school, marriage, and even a trip to a vacation destination. The cause of stress (i.e. the stressor) is attributed to for example the unfriendly boss, teacher or partner, or the computer that crashed right before a major deadline. Although it seems plausible and justifiable that people or objects are responsible for being stressed out, it mainly depends on how people experience and evaluate the stressor.

Whether a stressor is pleasant or threatening depends on the individual appraisal of the situation, which is based on previous experience and expectations of the outcome (Ursin & Eriksen, 2004). The appraisal process occurs in two stages. Primary appraisal refers to the way a person construes the significance of an encounter for his or her well-being, which can be irrelevant, harmless, harmful, threatening, or challenging, the latter three being forms of stress appraisal (Lazarus & DeLongis, 1983). In a secondary appraisal, various coping options are evaluated, such as altering the situation or accepting it (see Folkman, Lazarus, Dunkel-Schetter, DeLongis & Gruen, 1986). According to Ursin et al. (1978), coping is the trust in one's own abilities to perform, and it influences subjectively reported fear and the stress responses. So, when it is unlikely that the computer in the above example can be fixed before the major deadline; i.e. the person has negative expectancies of the outcome, then a stress experience and response is likely to occur. For seniors the encounter might even be more stressful than for young adults, because they have fewer means to cope with the demand.

On a psychological (see Karel, 1997), physiological (see McEwen & Lasley, 2002), and immunological level (see Hawkley & Cacioppo, 2004), seniors are more vulnerable to stress than are young adults. On the other hand, over the years seniors have found ways to cope with familiar stressors, and this might be an important factor why seniors might be more stressed when interacting with new media, because time to find ways to cope with new stressors is limited and costly. Within the computer domain, seniors generally are less experienced (see Brynin et al., 2004), and are less confident about their judgements (Marquié et al., 2002), than young adults. Within seniors, gender differences have been found in computer anxiety; female seniors report more computer anxiety than males (see Czaja et al., 2006; Karavidas et al., 2005). Although some studies have addressed stress in senior computer interaction (see Birdi & Zapf, 1997; Czaja & Sharit, 1993), little is known about the mechanisms behind stress in senior computer interaction. Two factors are considered here that could influence the stress experience and response in senior computer interaction (see figure 4.1).

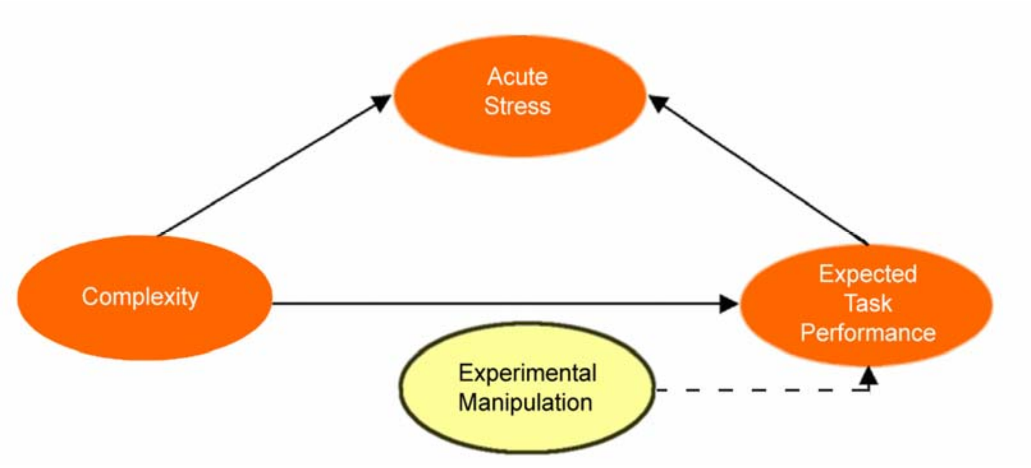

Figure 4.1: Theoretical model of acute stress in senior computer interaction

The complexity of the interface is the first factor that might have an influence on acute stress in senior computer interaction (see figure 4.1). The complexity of task performance is dependent on the difficulty of a task, on the level of experience of the user, but also on the tools with which (e.g. a retrieval system) a task can be performed. The two interfaces presented in figure 4.2 show what is meant by the complexity of a tool. The interface on the left was validated as a low complexity retrieval system for seniors, whereas the interface on the right was validated as a high complexity retrieval system for seniors (see chapter 2, this thesis). In that study, acute stress was observed (chapter 2.1.3, this thesis) when seniors performed search tasks on the high complexity system, while no stress was observed when seniors performed search tasks on the low complexity system (chapter 2.3.3, this thesis). The low complexity system is attuned to the knowledge and capabilities of seniors, by means of the native language, large fonts and buttons in the design. With a low complexity system, seniors are provided with means to cope with task complexity demands. It should be mentioned that no significant influence of complexity on task performance and subjective- and physiological stress was found in the experiment presented in chapter 3 (this thesis), because the experiment was very likely flawed by the automatic presentation of tasks.

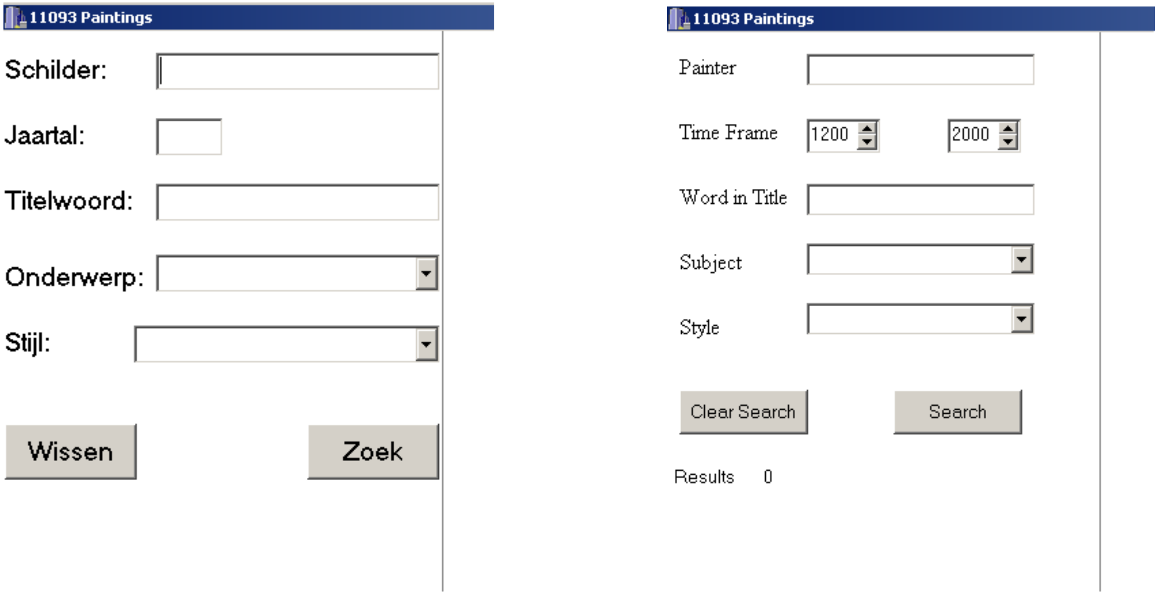

Figure 4.2: Low (left) and high (right) complexity art image retrieval system

Expected task performance is the second factor that might have an influence on stress in senior computer interaction (see figure 4.1). Expected task performance is a personal expectation of performance of the task at hand. According to the Cognitive Activation Theory of Stress (CATS) (Ursin & Eriksen, 2004), the stress response is dependent on the available resources, the affective value and the probability of the expected event. Stress is expected to be high when the affective value of the expected event is highly unattractive and when the perceived probability of the expected event is high, i.e. high-perceived probability of an unattractive event leads to high, rather than low arousal. If the outcome expectancies are low, i.e. a high chance of low performance, a stress experience and response is likely to happen.

Czaja et al. (2006) state that as computer self-efficacy is an important predictor of computer anxiety, it is critical to ensure that older people receive encouraging feedback during training, and experience some degree of success. Self-efficacy determines whether coping behaviour will be initiated, how much effort will be expended, and how long it will be sustained in the face of obstacles and aversive experiences (Bandura, 1977). Expected task performance is more in line with the 'outcome expectancies' of Ursin and Eriksen (2004), than with the self-efficacy concept of Bandura (1977). Self-efficacy is a broad concept, which is also related to other mechanisms than stress; e.g. learning, motivation, and addiction.

To reduce stress, positive outcome expectancies (coping) could be induced by positive performance feedback and feed-forward information (e.g. "You will do a good job!"). When coping is induced, trust is given in the user's own abilities to perform. Jussim, Yen and Aiello (1995) found support for the idea that people's self-perceptions are highly sensitive to both interpersonal evaluations, such as praise and criticism, and objective performance, such as test scores. Further, it was found that people felt better and expected to perform better after receiving positive feedback than after receiving negative feedback (Jussim et al., 1995).

Complexity will not only have an influence on stress in senior computer interaction, but also on the expectancies seniors have about their task performance (see figure 4.1). After successfully performing tasks on a system, the expected task performance people store in memory is expected to be positive. In addition, an interface that is compatible with the knowledge and capabilities of the user (i.e. a low complexity system), will increase task performance (see chapter 2, this thesis). Eventually, the complexity of an interface will be reduced after a number of successfully performed tasks, i.e. when the user becomes more skilled and experienced.

In addition to the two factors that could have an influence on stress, it is interesting to explore the relation between task performance and stress. From the 1960s the relation between arousal and performance has been described as an inverted U-curve (see Teigen, 1994), in which moderate levels of arousal will improve performance up to some maximum. In the Hebbian inverted U-curve, the optimum level of arousal is higher in a simple task than in a complex task (see Teigen, 1994). The inverted U-curve is the most well-known curve, but many different curves have been found over the years (see Teigen, 1994), so it is unclear what relation to expect between task performance and stress in senior computer interaction. But, since task complexity is dependent on the complexity of a system, it is expected that differences will be found in the relation between task performance and stress when system complexity is low or high. Aside from a possible relation between task performance and stress in senior computer interaction, it would also be interesting to explore the relation between task performance and expected task performance. It is expected that seniors' expectancies about future performance will be high when their task performance is high, and that their expectancies about future performance will be low when their task performance is low.

So far, no experimental data are available that indicate a (causal) relation between the complexity of an interface and stress in seniors (see Birdi & Zapf, 1997; Czaja & Sharit, 1993; Hamborg & Grief, 2003; Hawthorn, 2000). It could be possible to increase coping and reduce stress by providing seniors tools that are compatible with their knowledge and capabilities, and by providing seniors positive expectancies of the outcome. To reduce stress in senior computer interaction, expected task performance (outcome expectancies) could be induced by positive feedback- and feed-forward information of task performance. Further, stress could be lowered by reducing the complexity of the tools (e.g. interface and interaction devices). If it can be verified that (induced) expected task performance influences the stress experience and response in senior computer interaction, then possible mechanisms in senior computer interaction underlying stress will be revealed. In the following experiment knowledge will be gathered about the effect of complexity and (induced) expected task performance on acute stress in senior computer interaction, i.e. an empirical verification of the relation between the factors displayed in figure 4.1.

4.2 Experiment

4.2.1 Method

When stress experiences and responses are studied, it is important that external environmental effects on stress are reduced to a minimum. Activity theory (Nardi, 1995) states that actions and objects have different meanings in different contexts. When possible and appropriate, scientists should avoid shaping the meaning of actions and objects by observing and measuring behaviour in a 'synthetic' laboratory setting. For this reason, the experiment was conducted in a controlled home environment (‘huiskamer’, shared living room) of the seniors (see figure 4.3), rather than in the controlled laboratory setting.

Figure 4.3: Participant in home environment

Ethics

Ethics should be considered in all research involving humans and animals as subjects. For this experiment, special attention was paid to ethical issues, because the subjects run the risk of experiencing acute stress. An ethicist commented upon the initial research protocol, and as a result, minor revisions were made to the protocol to meet current ethical research standards, e.g. by means of the informed consent the participants were notified that they could experience stress during the experiment.

Participants

Thirty seniors participated at the Grevelingen Care Centre in 's-Hertogenbosch (The Netherlands), 13 male and 17 female (mean age = 71.04, SD = 5.35), who were paid €10,- (≈ $14,-). Criteria for participation were: familiarity with the English language, subjectively healthy, and having intermediate experience with PCs (beside Microsoft Word, they also used other programs). To control for variables that influence the accuracy of the physiological measurements (see O'Brien et al., 2003), participants were asked if they had diabetes, if they used blood pressure (BP) medication, heart medication, or both.

In total, fifty-five seniors signed up for participation from which eight cancelled participation, eleven seniors were not able to perform a preliminary test (see procedure), and six seniors needed two hours to complete one of two conditions. Ten of the eleven seniors that were not able to perform the preliminary test did not have experience with PCs. Four of these stated that they did not want to use PCs and three stated that they would like to use PCs, but found it too difficult. One stated that she did not have experience with PCs because of her age, and another one questioned if it would still be possible to learn to work with PCs.

Material

The experiment was conducted on a Dell latitude D800 laptop with a Pentium M processor and with a 15.4" wide screen (1280x800-pixel resolution). The low complexity system and the high complexity system (see figure 4.2) were used as the platforms on which the participants performed search tasks. The design and functionality of the two complexity systems are described in more detail in chapter 2 (this thesis). BP was measured with an arm cuff connected to an automatic BP measurement device (Omron M6), using the oscillometric method, and is clinically validated according to the International Protocol of the European Society of Hypertension (Topouchian et al., 2006). A continuous heart rate (HR) recorder at 1024Hz (TMSi Mobi) connected to a second laptop, was used to measure HR by means of an ear clip. The tasks were written on paper. A stopwatch was used to measure task time. All instructions were read from pre made printed cards.

Tasks

Two matching sets of four practice search tasks were to be completed on the two search interfaces. The practice search tasks served to make it possible to give feedback- and feed-forward information about task performance. The search tasks concerned the name of a painter (single task) and the title of a painting plus three other paintings made by the same painter (compound task) two times each. A single task implied using the search interface once, for example: "Find a painting of a dog". A compound task implied using the search interface twice: "Who painted 'Coymans' and who painted 'Ramparts'?". After a participant filled in or selected a search word on the image retrieval system (e.g. ‘Coymans’ at the ‘Titelwoord’ or ‘Word in Title’ fill-in box, see figure 4.2), the search button could be pressed, and then matching art images were presented on screen, on the right of the search interface. To start with a new single search task, or the second search task in a compound task, ‘Wissen’ or ‘Clear Search’ could be pressed, which cleared all the fill-in boxes and the art images on the right of the search interface.

Two matching sets of twelve search tasks had to be completed on the two systems. The twelve search tasks served to compare the task performance of the seniors on the low- and high complexity systems, and to examine the effect of induced expected task performance and complexity on stress and expected task performance over time. The search tasks in each set consisted of six single and six compound tasks. The single tasks concerned the title of a painting, the subject of a painting, the year of creation, the style of a painting, and two replications of painter name and subject. The compound tasks concerned the title of a painting plus two other paintings made by the same painter, the year of creation of two paintings, the style of two paintings, the number of paintings available of two painters, the subject of two paintings, and the titles of two paintings.

Procedure

The participants received by mail an informed consent form and a questionnaire to collect the participants' characteristics. On arrival, the signed informed consent form and questionnaire were registered and after this, the participants performed a preliminary test in which they had to type their own name in a box and had to press a button on the screen with the mouse. This test was introduced to make sure that all participants had basic knowledge of computer input devices. Information was given about the experimental procedure and it was explicitly stated that they were allowed to stop the experiment whenever they wanted. They had five minutes to relax, BP was measured twice and subjective stress was measured (pre-measurement). After this the participants were given a short introduction about the search tasks they were to perform, their expected task performance was measured (pre-measurement), this was followed by the four practice search tasks. Two of the practice tasks were single and two were compound, both tasks had a time limit of one minute. Negative or positive feedback was given about their performance (see manipulations section and figure 4.5), their expected task performance was measured again (post1-measurement), and after that their BP and subjective stress was measured (post1-measurement). Next, participants carried out the twelve search tasks. Time limits were one minute for the single tasks and two minutes for the compound tasks. This was followed by measurement of BP and subjective stress (post2-measurement). After this expected task performance was measured (post2-measurement). The participants had a five-minute break to relax (everybody wanted to proceed without a break) and after that, the procedure was repeated for the other system. The post2-measurement of acute stress of the first (preceding) condition was used as the pre-measurement of the second (following) condition (see figure 4.4 for an overview of the procedure). In the debriefing, any possible emotional distress during the experiment was reduced by restorative communication. Information was given about the nature and goal of the research. The performance feedback screen was shown again and it was explicitly stated that the feedback given was fictitious. The total experiment took about one hour and thirty minutes per participant.

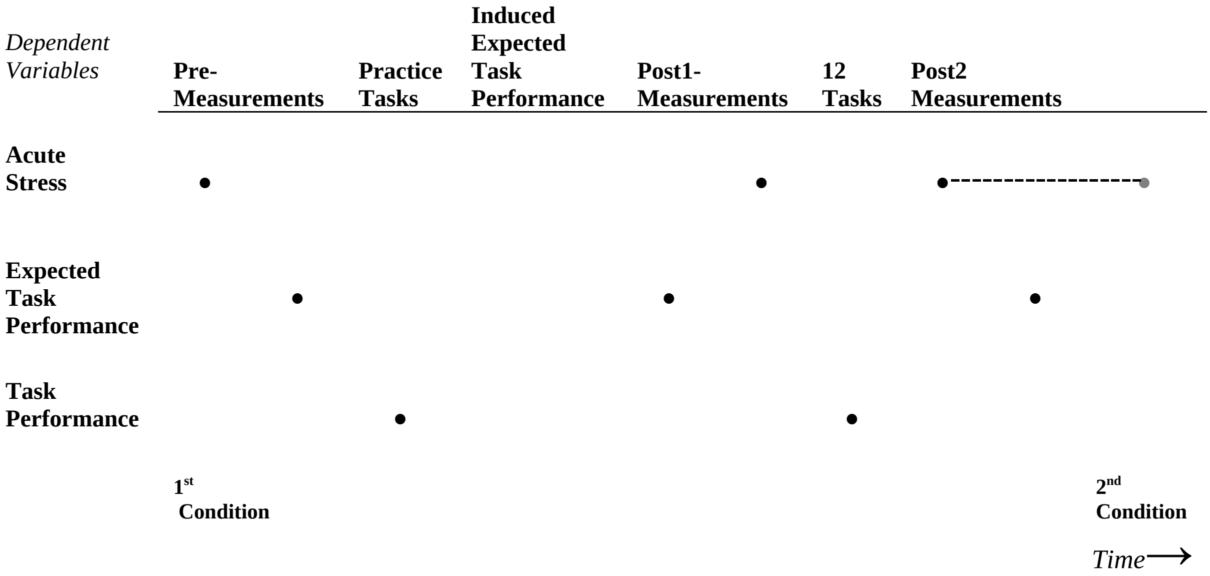

Figure 4.4: Procedure timeline

Dependent variables

Acute stress

Subjective stress was measured by having participants fill in a rating scale on four attributes: stressed (gestressed), tensed (gespannen), irritated (geprikkeld), and nervous (zenuwachtig). These attributes had to be rated on an 11-point rating scale (0 = not at all, 10 = totally). The mean Cronbach's alpha of the scale was 0.81 in the experiment presented in chapter 3 (this thesis).

Physiological stress was measured by systolic- and diastolic BP, and HR. The mean value of two readings was used for analysis. HR was measured at the same time BP was measured, and continuously at a 1024 sampling rate. By measuring the HR continuously at a large frequency, Heart Rate Variability (HRV) could be calculated. Low HRV is an indicator of high arousal (e.g. stress and mental load), while high HRV is an indicator of low arousal (e.g. relaxation).

Expected task performance

Before the practice tasks (pre), expected task performance was measured by the question: "Now you will receive four practice tasks that are executed with a computer program with which you can search for images. How likely do you think you will perform the tasks correctly?", that had to be rated on an 11-point scale ranging from 0 (very unlikely) to a 100 (very likely) with 10 unit intervals. This type of scale has also been used to measure self-efficacy (Wiedenfeld, O'Leary, Bandura, Brown, Levine & Raska, 1990). The scale reflects the subjective probability that a set value can be met with the available resources.

After the practice tasks (post1), expected task performance was measured by the question: "Now you will receive twelve tasks that are executed on the computer program with which you just searched for images. How likely do you think you will perform the tasks correctly?", that also had to be rated on the 11-point scale.

After the twelve search tasks (post2), expected task performance was measured by the question: "If you would work with this computer program in the future, how likely do you think you could perform the tasks correctly?", that also had to be rated on the 11-point scale.

Task performance

Task performance was used to check whether the low complexity system is more effective and efficient (ISO 9421-11) for retrieving digital art images than the high complexity system. Effectiveness was measured by the number of search tasks completed correctly, and efficiency was measured by total task time. When a participant was unable to perform a task within the time limit of the specific task, one minute for the single task and two minutes for the compound task, the time limit was coded as the efficiency score. Satisfaction (ISO 9421-11, 1998) was not measured due to limited available time.

Manipulations

Complexity was induced by a low- and high complexity art image retrieval system (see figure 4.2), that have been validated, respectively, as a high- and low accessible system for seniors (chapter 2, this thesis).

Expected task performance was induced by fictitious feedback-, comparison- and feed-forward information. To induce low expected task performance, a performance graph was shown (see figure 4.5) and the participants were told by the experimenter: "The green line represents the score of 114 other persons and the red line represents your score. You did not perform well. You performed worse than the other persons did." A button labelled 'Calculate Prognosis' was pressed by the experimenter and the participants were told: "I do not expect you to perform well." To induce high expected task performance, a performance graph was shown and the participants were told: "The green line represents your score and the red line represents the score of 114 other persons. You performed well. You performed better than the other persons did." The button 'Calculate Prognosis' was pressed and the participants were told: "I expect you to perform well."

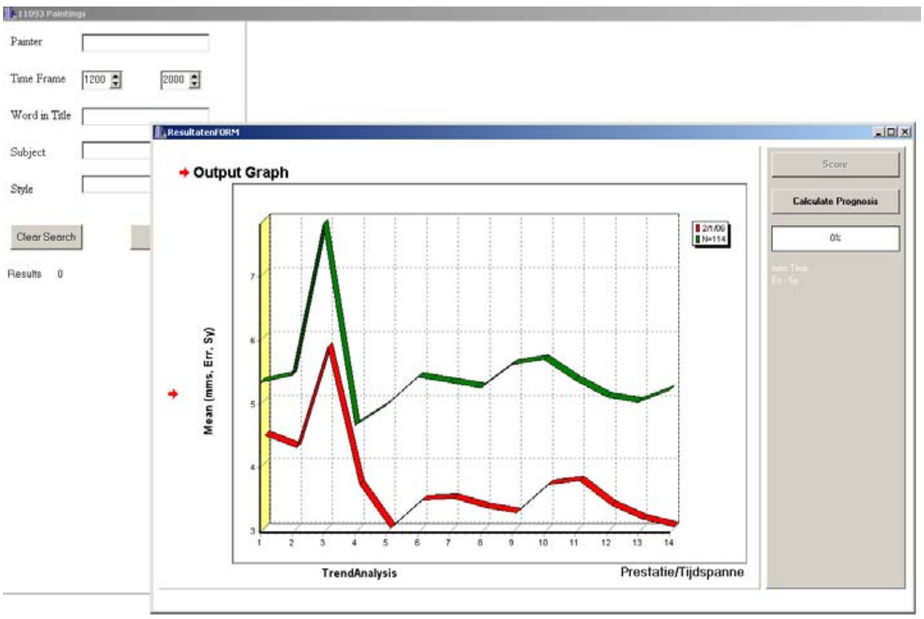

Figure 4.5: Negative performance outcome graph (upper line is green, bottom line is red)

Design

A within-subjects design was used. The participants performed two matched sets of similar tasks on the two systems. The presentation order of the two task sets and the two systems were counterbalanced.

There were four feedback groups: participants that received positive feedback twice, participants that received negative feedback twice, participants that received positive feedback first and negative feedback second, and participants that received negative feedback first and positive feedback second.

4.2.2 Results

Stress scale reliability

The mean Cronbach's alpha of the five repetitions of the subjective stress measurement scales was 0.86 (range 0.82-0.88); this is high enough to use the mean value of the four items as a subjective stress variable.

Task performance by task sets

Because two different, but matching, task sets were used, the task performance with both task sets should be similar. If no difference is found, task set can be disregarded as a between-subjects factor in the following analyses. A paired samples t-test was performed to compare the two task sets by the number of tasks performed correctly and the total time it took to perform the tasks. No significant difference was found between the task sets on the number of tasks performed correctly [$t(29) = .392$, $p = .698$], and no significant difference was found between the task sets on the total time it took to perform the tasks [$t(29) = .734$, $p = .469$]. See table 4.1 for means and standard deviations.

Table 4.1: Averages (*M*) and standard deviations (*SD*) of task performance on task set 1 and task set 2

		Task Set	
		1	**2**
Number of Tasks correct	*M*	5.27	4.83
(max.=12)	*SD*	3.28	3.26
Task Time	*M*	699.7	683.0
(in seconds, max.=1080)	*SD*	213.7	234.5

The effect of temporal order on task performance, expected task performance and subjective stress

It is interesting to examine the condition order effect on the dependent variables, because it can shed light on a possible temporal effect. It was expected that when seniors become more experienced over time that their task performance should be higher in the following condition compared to the preceding condition.

Furthermore, the seniors' expected task performance could increase and their subjective stress could decrease over time.

A paired samples t-test was performed to compare the task performance of the preceding and following condition by the number of tasks performed correctly, and the total time it took to perform the tasks (see table 4.2 for averages and standard deviations). Significantly more tasks were performed correctly in the following condition compared to the preceding condition [$t(29) = 2.688$, $p = .012$]. The participants were significantly faster in performing the search tasks in the following condition compared to the preceding condition [$t(29) = 3.178$, $p = .004$].

Table 4.2: Mean task performance (*M*) and their standard deviations (*SD*) for the preceding and following conditions

		Condition	
		preceding	**Following**
Number of Tasks correct	*M*	4.33	5.77
(max.=12)	*SD*	3.16	3.23
Task Time	*M*	747.7	635.0
(in seconds, max.=1080)	*SD*	202.4	230.8

A repeated-measures ANOVA was performed with time (3 levels: pre, post1 and post2) as a within-subject factor, with expected task performance and subjective stress as the dependent variables. Condition (preceding and following) was used as a between-subject factor.

No significant main effect of condition was found on expected task performance [$F(1, 58) = .527$, $p = .471$]. No significant main effect of condition on subjective stress was found either [$F(1, 58) = .000$, $p = 1.000$]. Table 4.3 shows averages and standard deviations.

A significant interaction effect was found between time and condition on expected task performance [$F(1, 58) = 12.8$, $p = .001$], with an effect size (partial η^2) of .181 and an observed power of .941. In addition, a significant interaction effect was found between time and condition on subjective stress [$F(1, 58) = 5.25$, $p = .026$], with an effect size (partial η^2) of .124 and an observed power of .805.

Table 4.3. Averages (*M*) and standard deviations (*SD*) of expected task performance and subjective stress per condition (preceding and following) and time (pre, post1, and post2)

		Expected Task Performance			Subjective Stress		
Condition		**Pre**	**Post1**	**Post2**	**Pre**	**Post1**	**Post2**
Preceding	*M*	55.67	34.67	47.00	2.03	2.89	2.73
	SD	18.88	23.89	27.43	1.82	2.23	2.20
Following	*M*	43.00	47.00	57.00	2.73	2.46	2.46
	SD	20.37	19.15	19.15	2.20	2.02	2.01

4.2.3 Manipulation checks

Complexity

A paired samples t-test was performed to compare the participants' task performance on the four practice tasks when complexity was low or high. No significant difference was found between the number of practice tasks performed correctly on the low- and high complexity systems [$t(29) = .250$, $p = .804$], and between the total time it took to perform the practice tasks [$t(29) = .331$, $p = .743$]. Table 4.4 shows means and standard deviations. No task performance difference was found on the practice tasks between the low- and high complexity systems, so it is unlikely that the expected task performance manipulation was influenced favouring one of the complexities.

Table 4.4: Averages (*M*) and their standard deviations (*SD*) of task performance on the practice tasks on the low- and high complexity systems

		Complexity	
		Low	**High**
Number of Tasks correct *(max.=4)*	*M*	1.30	1.23
	SD	1.02	0.90
Task Time *(in seconds, max.=240)*	*M*	189.53	193.00
	SD	39.45	35.51

A paired samples t-test was performed to compare the complexity of the two interfaces that were used, by the number of tasks performed correctly, and the total time it took to perform the tasks (see table 4.5 for averages and standard deviations). Significantly more tasks correct were performed when complexity was low compared to high [$t(29) = 2.672$, $p = .012$]. The participants were significantly faster in performing the search tasks when complexity was low compared to high [$t(29) = 3.184$, $p = .003$].

Table 4.5: Averages (*M*) and standard deviations (*SD*) of task performance on the low- and high complexity systems

		Complexity	
		Low	**High**
Number of Tasks correct	*M*	5.87	4.23
(max.=12)	*SD*	3.59	2.69
Task Time	*M*	642.07	740.67
(in seconds, max.=1080)	*SD*	237.69	198.17

Induced expected task performance

A repeated-measures analysis (ANOVA) was performed with time (2 levels: pre and post1) as a within-subject factor, with expected task performance as the dependent variable. Induced expected task performance (positive and negative feedback) was used as a between-subject factor.

A significant interaction effect was found between time and feedback on expected task performance [$F(1, 58) = 24.6$, $p = .000$], with an effect size (partial η^2) of .298 and an observed power of .998. After the participants received negative feedback their expected task performance dropped considerably, while after positive feedback there was a small gain in expected task performance (see table 4.6 for averages and standard deviations).

Table 4.6: Averages (*M*) and standard deviations (*SD*) of expected task performance per induced expected task performance (positive and negative) and time (pre and post1)

Induced Expected Task Performance		Expected Task Performance	
		Pre	**Post1**
Positive	*M*	46.13	51.29
	SD	23.05	18.21
Negative	*M*	52.76	29.66
	SD	17.09	21.13

In the following analyses, the physiological stress data is presented first, without the subjective stress data and expected task performance data. This was done because a third of the participants had to be excluded from the sample.

4.2.4 The effect of complexity and feedback on acute physiological stress

Ten of the thirty participants were excluded in the following analysis, because eight used BP medication, heart medication, or both, from which two also had diabetes Type 2, and one solely had diabetes Type 2, which could have influenced the BP and HR readings during the experiment. From one of the participants the data were missing because of a technical failure of the BP measurement device. The continuous HR data have not been analyzed, due to erroneous log files.

A repeated-measures ANOVA was performed with complexity (2 levels: high and low) and time (3 levels: pre, post1 and post2) as within-subject factors, with systolic BP, diastolic BP and HR as the dependent variables. Feedback group (positive-positive, negative-negative, positive-negative, negative-positive) was used as a between-subject factor.

Complexity

No significant main effect was found of complexity on systolic BP, diastolic BP, and HR. A significant main effect was found of time on systolic BP [$F(1, 16) = 4.74$, $p = .045$], with an effect size (partial η^2) of .229 and an observed power of .534. The participants' systolic BP was higher before expected task performance was induced (pre) than after it was induced (post1). Furthermore, a significant main effect was found of time on HR [$F(1, 16) = 29.6$, $p < .001$], with an effect size (partial η^2) of .649 and an observed power of .999. The participants' HR was higher before expected task performance was induced (pre) than after it was induced (post1), and HR was higher after expected task performance was induced (post1) than after the twelve search tasks were performed (post2).

Induced expected task performance

To analyze the effect of induced expected task performance (positive or negative feedback) on physiological stress, a repeated-measures ANOVA was performed with time (3 levels: pre, post1 and post2) as a within-subject factor, with systolic BP, diastolic BP, and HR as the dependent variables. Induced expected task performance (positive and negative feedback) was used as a between-subject factor.

No significant main effect was found of feedback on systolic BP [$F(1, 38) = .096$, $p = .759$], on diastolic BP [$F(1, 38) = .010$, $p = .923$], and on HR [$F(1, 38) = 1.35$, $p = .252$].

No significant interaction was found between time and feedback on systolic BP [$F(1, 38) = .085$, $p = .772$], on diastolic BP [$F(1, 38) = .370$, $p = .547$], and on HR [$F(1, 38) = 1.60$, $p = .214$].

4.2.5 The effect of complexity and feedback on subjective stress and expected task performance

A repeated-measures ANOVA was performed with complexity (2 levels: high and low) and time (3 levels: pre, post1 and post2) as within-subject factors, with subjective stress and expected task performance as the dependent variables. Feedback group (positive-positive, negative-negative, positive-negative, negative-positive) was used as a between-subject factor.

No significant main effect was found of complexity on subjective stress [$F(1, 26) = .465$, $p = .502$], with an effect size (partial η^2) of .018 and an observed power of .101 Participants did not experience more stress when complexity was high than when complexity was low (see table 4.7 for averages and standard deviations).

Table 4.7: Averages (*M*) and standard deviations (*SD*) of subjective stress per level of complexity (low and high) and time (pre, post1, and post2)

		Subjective Stress		
Complexity		**Pre**	**Post1**	**Post2**
Low	*M*	2.30	2.73	2.48
	SD	2.06	2.10	1.99
High	*M*	2.45	2.62	2.70
	SD	2.04	2.16	2.23

A significant main effect was found of feedback group on subjective stress [$F(3, 26) = 4.84$, $p = .008$], with an effect size (partial η^2) of .358 and an observed power of .855. Post-hoc tests with a Sidak correction indicated that the subjective stress of the participants in the pos-neg feedback group was significantly higher than the subjective stress of the participants in the pos-pos feedback group [$t(26) = 3.234$, $p = .020$].

A significant main effect was found of complexity on expected task performance [$F(1, 26) = 11.3$, $p = .002$], with an effect size (partial η^2) of .303 and an observed power of .899. Participants expected to perform better when complexity was low than when complexity was high (see table 4.8 for averages and standard deviations).

Table 4.8: Averages (*M*) and standard deviations (*SD*) of expected task performance per level complexity (Low and High) and time (pre, post1, and post2)

		Expected Task Performance		
Complexity		**Pre**	**Post1**	**Post2**
Low	*M*	49.67	45.00	58.67
	SD	22.36	23.31	20.13
High	*M*	49.00	36.67	45.33
	SD	18.82	20.89	25.96

An additional analysis was performed to examine the effect of induced expected task performance (positive and negative feedback) on subjective stress before and after the feedback was given.

A repeated-measures ANOVA was performed with time (2 levels: pre and post1) as a within-subject factor, with subjective stress as the dependent variable. Induced expected task performance (positive and negative feedback) was used as a between-subject factor.

A significant interaction effect was found between time and feedback on subjective stress [$F(1, 58) = 7.23$, $p = .009$], with an effect size (partial η^2) of .111 and an observed power of .753. Participants' subjective stress increased after they received negative feedback and decreased after they received positive feedback (see table 4.9 for averages and standard deviations).

Table 4.9: Averages (*M*) and standard deviations (*SD*) of subjective stress per induced expected task performance (positive and negative) and time (pre and post1)

Induced Expected Task Performance		Subjective Stress	
		Pre	**Post1**
Positive	*M*	2.52	2.34
	SD	2.24	2.05
Negative	*M*	2.22	3.04
	SD	1.82	2.17

4.2.6 Correlations of post2 subjective stress and post2 expected task performance with task performance

Correlations were calculated to explore if there were relations between total task times (efficiency) and total tasks correct (effectiveness) and the stress experience and task performance expectancies of seniors after the tasks were performed, and if so, what type of relations exist. These correlations were calculated for all the data, and separately for the low- and high complexity systems.

Bivariate correlations indicated a significant Pearson correlation at the 0.01 level (2-tailed) between efficiency and post2 subjective stress [r = .363], and between effectiveness and post2 subjective stress [r = -.358] (see figure 4.6 & 4.7 for scatter plots). In addition, a significant correlation was found at the 0.01 level between efficiency and post2 expected task performance [r = .381], and a significant correlation was found at the 0.05 level between effectiveness and post2 expected task performance [r = -.286]. Subjective stress increased with increasing total task times and decreasing total tasks correct, furthermore, expected task performance increased with decreasing total task times and increasing total tasks correct.

When the correlations were calculated separately for the low- and high complexity systems, a significant Pearson correlation was found only for low complexity at the 0.05 level (2-tailed) between efficiency and post2 subjective stress [r = .383], and a significant correlation was found only for high complexity between effectiveness and post2 subjective stress [r = -.436]. Although not significant, similar correlation strengths were found between efficiency and post2 subjective stress for high complexity [r = .345], and between effectiveness and post2 subjective stress for low complexity [r = -.304]. In addition, a significant correlation was found only for low complexity at the 0.05 level (2-tailed) between efficiency and expected task performance [r = -.391], and between effectiveness and expected task performance [r = .446]. When complexity was low, expected task performance increased with decreasing total tasks times and increasing total tasks correct.

From figure 4.6 and 4.7, it appears that high subjective stress is mainly paired with high total task times and low total tasks correct. So why do some seniors report low stress while task performance is low? To come up with a possible explanation, it is interesting to look into the characteristics of, and the type of manipulations the seniors received that performed poorly but reported low stress. The ages, genders, and computer experiences of these seniors were checked and the complexity and induced expected task performance they received. In addition, temporal order was checked, i.e. if a data point belongs to the preceding or following condition. For induced expected task performance, complexity, and temporal order, substantial deviations from the means of these characteristics and manipulations were found. The scatter plots in figure 4.6 and 4.7 can be divided into four quadrants, centred on mean post2 subjective stress (*M* = 2.59), and on mean total task times (*M* = 691.27) and on mean total tasks correct (*M* = 5.05). Of the data points (*n* = 15), 66.7 % in the upper left quarter of figure 4.6 fall into the positive induced expected task performance- and high complexity group, and 60% fall into the preceding condition. Of the data points (*n* = 13) in the lower left quadrant of figure 4.7, 76.9% fall into the positive induced expected task performance group. Other interesting deviations from the means were found in the lower right quadrant of figure 4.6; 62.5 % negative induced expected task performance and low complexity (*n* = 8), and in the lower left quadrant; 65% low complexity and 60% following condition (*n* = 20). In addition, 71.4% of the data points (*n* = 7) in the upper right quadrant of figure 4.7 fall into the negative induced expected task performance group.

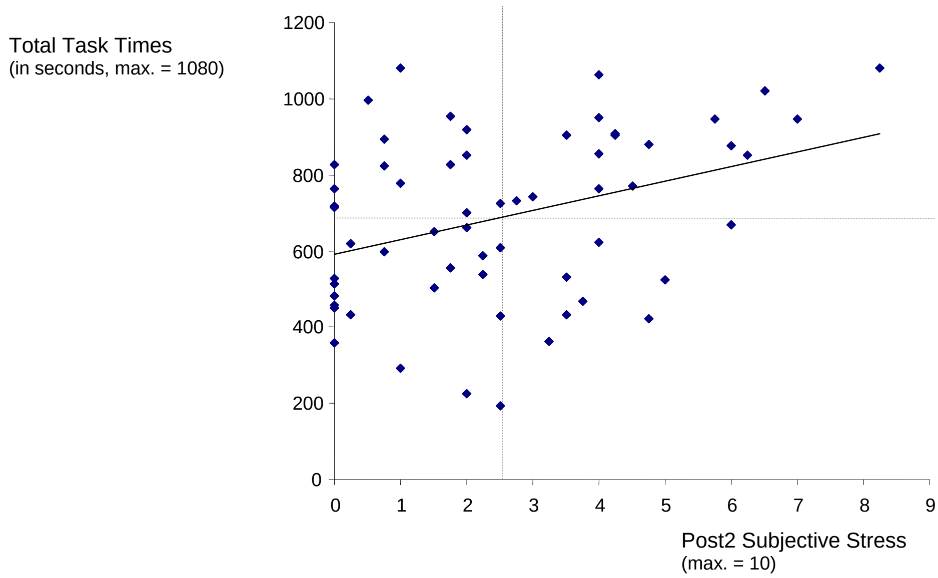

Figure 4.6: Scatter plot of total task times by post2 subjective stress, with trendline and quadrants centred on mean post2 subjective stress (*M* = 2.59) and on mean total task times (*M* = 691.27).

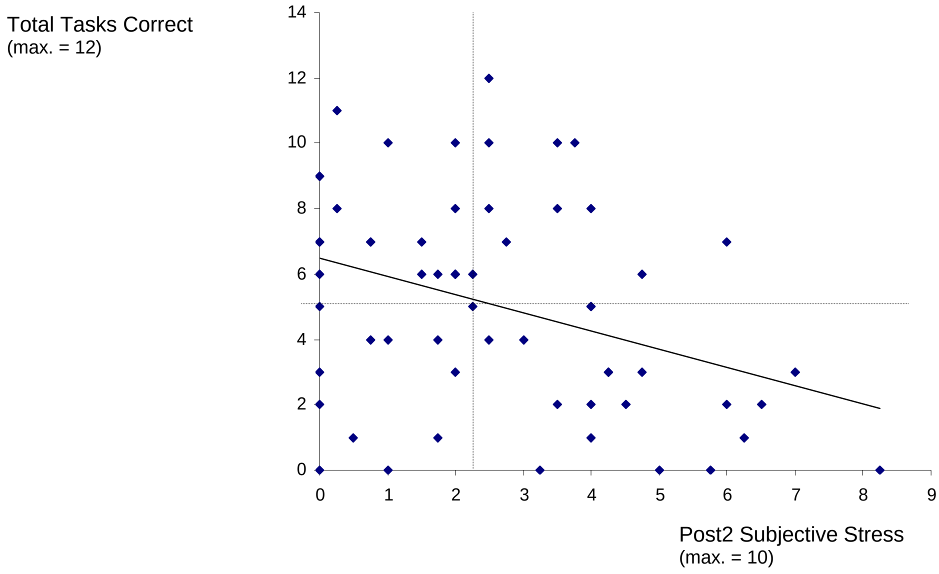

Figure 4.7: Scatter plot of total tasks correct by post2 subjective, with trendline and quadrants centred on mean post2 subjective stress ($M = 2.59$), and on mean total tasks correct ($M = 5.05$).

4.3 Discussion

The experiment presented in this chapter was conducted to examine the effect of complexity and (induced) expected task performance on acute stress in senior computer interaction (see figure 4.8).

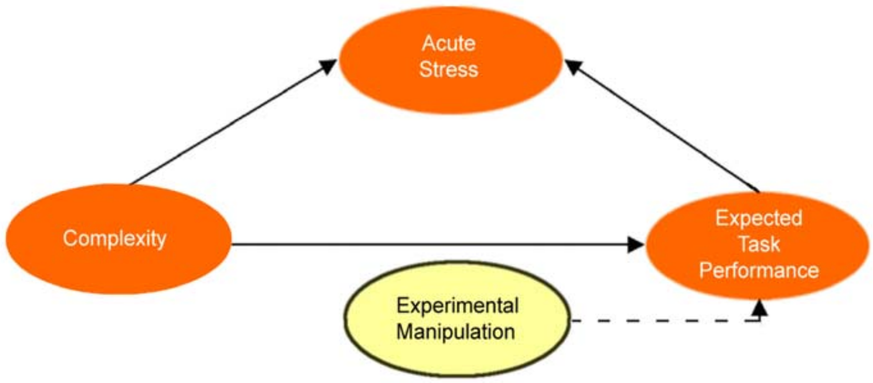

Figure 4.8: Theoretical model of acute stress in senior computer interaction

Complexity did not have a significant influence on acute subjective- and physiological stress in senior computer interaction. The stress experience and stress response of seniors were not significantly higher when they performed tasks on a highly complex system than on a low complexity system.

Induced expected task performance did significantly influence acute subjective stress in senior computer interaction. Subjective stress increased after the seniors received negative feedback- and feed-forward information, and decreased when they received positive feedback- and feed-forward information. The effect of induced expected task performance on expected task performance was quite large. Seniors' expected task performance was higher after they received positive feedback- and feed-forward information than when they received negative feedback- and feed-forward information.

Complexity did have a significant effect on the seniors' expected task performance. Seniors expected to perform better when they performed search tasks on a low complexity system than when they performed tasks on a high complexity system. In line with the findings presented in chapter 2 (this thesis), seniors were more effective and efficient in performing the search tasks on the low complexity system than on the high complexity system. The low complexity system provides intermediate computer experienced seniors with high accessibility. All seniors preferred using the low complexity system and many wanted to take a copy at home or said, "When will this system be available on the market? I will certainly buy it."

There was a clear effect of order. In the preceding condition, seniors' expected task performance decreased and subjective stress increased after feedback was given, while in the following condition, there was a slight increase in expected task performance and decrease in subjective stress. The order effect might have been caused by a gain in confidence about future performance, because the seniors were more effective and efficient in performing the search tasks in the following condition compared to the preceding condition. Stress levels were reduced and expected task performance is increased, probably because uncertainty is reduced. In short, a learning or carry-over effect was present. This might have been reduced by a large time interval, maybe up to one week, between the two conditions.

Although complexity has no effect on subjective stress, task performance is related to subjective stress. An inverted U-curve was not found, but a linear relation was found between subjective stress and task performance. Subjective stress increases with decreasing task performance. It should be mentioned that high subjective stress is mainly paired with low efficiency and effectiveness, while low subjective stress is paired with low to high efficiency and effectiveness. A possible cause for seniors to report low stress while their performance was low was found in the feedback they received. Most of these seniors received positive induced expectancies about their performance, which likely have influenced them to report low stress while their task performance was low. This is in line with the significant influence that was found of the expected task performance manipulation on acute subjective stress and the findings of Ursin et al. (1978), who found that the subjective feeling of being able to perform reduces the stress responses and not the actual performance. In addition to the linear relation between subjective stress and task performance, a linear relation was also found between expected task performance and task performance. Expected task performance increases with increasing task performance. When the correlations were calculated separately for low- and high complexity, a significant linear relation was found only for low complexity between expected task performance and task performance. Because complexity had a significant influence on efficiency, effectiveness and expected task performance it is quite plausible to find that high expected task performance is paired with high efficiency and effectiveness, and that low expected task performance is paired with low efficiency and effectiveness. During the debriefing, some of the participants could not be convinced that it was not their performance that was shown but a performance that was made up (see figure 4.5). It could be argued whether it is ethical to give people false feedback about their performance. Misinformation (deception) was used to change the task outcome expectancies of the seniors, and in addition to change their level of subjective stress. Because deception was used, the technology that presented the feedback information falls outside the realm of persuasive technology (see IJsselsteijn, De Kort, Midden, Eggen & Van den Hoven, 2006). Persuasive technology is defined as a class of technologies that are intentionally designed to change a person's attitude or behaviour, and persuasion implies a voluntary change of behaviour or attitude or both. Atkinson (2006) points out that persuasion that operates without the user being aware of the programmers' intent might be ethical if the change in attitude, behaviour or belief is motivated from the perspective of wisdom, benevolence and genuine care for others. Positive deceptive feedback information could fall into this type of persuasion, because it is favouring the users' self-efficacy and well-being. The technology that was used in the experiment to give false positive feedback can be easily amended so that the user can be in control of false positive feedback information, e.g. can voluntary turn it on or off. With such an adjustment and positive attitudinal changes in mind, positive deceptive technology could be a part of persuasive technology.

It should be noted that eleven seniors could not participate because they were unable to perform the preliminary test. They were not able to type their own name in a text box, use the mouse to click a button, or both. About 20% of the seniors that signed up for participation did not have the basic computer input skills to perform a simple task, and this shows that common computer input devices are complex and provide few affordances to a considerable number of seniors. Some of the seniors mentioned that the location of the characters on the keyboard were odd, they preferred an alphabetical layout. The alphabetical (ABCDE) keyboard of John Parkinson (CNET, 2006) might be of help for seniors with little computer experience. Others used their index finger to click the button on the screen rather than on the mouse; a touchscreen would easily overcome this type of input uncertainty.

Throughout the five-week stay at the care centre, we talked with at least sixty residents and it became apparent that many believe that they do not count as full members of society. The self-efficacy of many residents was low and almost all of them referred to the things they achieved when they were young. It should be mentioned that most of these residents were healthy and active persons with little or no signs of age-related cognitive impairments. The study presented in this chapter shows that the expectancies seniors have about how they will perform on a computer system decreases when they receive negative feedback and feed-forward information about their performance. This finding could have several practical implications for system designers, computer course teachers and senior caregivers.

System designers could implement affective aids in the form of positive feedback messages in (senior) computer software, or at least, try to avoid negative feedback messages. Similar to a virtual coach that motivates athletes to exercise behaviour (see Eyck, Geerlings, Karimova, Meerbeek, Wang, IJsselsteijn, De Kort, Roersma & Westerink, 2006), seniors' expected task performance could be increased by a virtual coach or teacher during computer interaction. Furthermore, it seems important to support memorability. Many of the seniors who participated in the experiment said that they forgot how to work with certain programs learned during computer courses, simply because they only used it a couple of times a year.

Computer course teachers could give special attention to seniors with low expected task performance. Negative feedback should be avoided and seniors should be focused on the tasks they are capable of performing. During the stay at the care centre, it became apparent that most inexperienced seniors would like to learn basic text processing, internet and e-mail skills. These seniors wanted to know how to do online banking, how to visit the web site of a grandchild or how to open an e-mail attachment. Learning how Solitaire works or how to insert a table in a document is less important to them. After the basic goals and needs of computer use are met, advanced skills can be introduced. The cost of learning new media is higher in most seniors compared to their younger counterparts, and the benefits should be high, clearly and apparent stated (Melenhorst, 2002). For example, if the main reason for teaching Solitaire during a computer course is to learn basic skills, then seniors should be informed that these costs are necessary before learning other programs.

Caregivers, who work with seniors, could focus even more on telling senior care receivers that they are capable of learning and using new media. At least, they should not be told that they are not capable of learning and using new media. In addition, to take this a step further, family and friends could be informed about the importance of this feed-forward information.

The ultimate goal is to provide users with positive outcome expectancies, and develop systems that increase coping and the level of expected task performance. Induced positive outcome expectancies might encourage seniors to interact with new media, and it will show them benefits in terms of higher self-efficacy. In addition, complexity has an effect on expected task performance, in other words: when the costs are reduced by a low complexity system, trust in the ability to perform is increased. Low complexity systems, specifically designed and engineered for seniors, are not guaranteed to be purchased by the intended user group. The Philips Easy Line products, specifically designed for seniors, were a market failure (Van Kuijk, 2005). On the other hand, the simPC, a personal computer specifically designed for seniors, seems to be a market success (simPC, 2006).

In conclusion, stress plays a role in senior computer interaction, and expected task performance even more. When seniors receive negative feedback about their performance, subjective stress increases and expected task performance decreases. System complexity has no influence on stress in senior computer interaction, but complexity does have an influence on expected task performance. The expected task performance of seniors decreases when they interact with a high complexity system.

A considerable number of seniors could not operate the mouse and the QWERTY keyboard. In future research, other input devices (e.g. a touchscreen or ABCDE keyboard) could be evaluated to increase accessibility. In addition, an information domain that is in general more personally related and relevant to seniors (e.g. the health domain instead of the art domain) could be used in a similar experimental setting to increase the social relevance of the developed information retrieval system. Moreover, additional validation within a different information domain will strengthen the theoretical model of stress in senior computer interaction. In the future, affective aiding could be developed and evaluated to lower stress and higher expected task performance in senior computer interaction.

Acknowledgements

Thanks to R. Lionarons, M.D. – Cardiology, for his advice considering BP measurement methods and the personnel and inhabitants of the Grevelingen for their hospitality.

Chapter 5: Understanding stress in senior computer interaction within the Alzheimer domain

5.1 Acute stress and information accessibility within the health domain

Increasing numbers of seniors search for health care related information on the Internet. In the United States of America there was a growth rate from the year 2000 to 2003 of 25% (Fox, 2004), 53% of seniors online have sought health information (Fox, 2001). The most popular health topic among online seniors is the information about a specific disease or medical problem (Fox, 2004). It is important that when seniors search for information about specific diseases they are concerned about already, they should not experience additional stress and negative feelings. Furthermore, it is important to provide high accessibility when seniors go online to search for health information.

The research presented in the previous chapters has all been conducted within the art domain, and the environments that were used to perform search tasks were art image retrieval systems. As mentioned in chapter 2 (this thesis), the art domain is recognizable and pleasurable for many people, hence it is unlikely to elicit negative feelings that could influence the interaction performance. In the research presented here, a domain is chosen that is more personally and emotionally related to seniors: the Alzheimer domain. In the Netherlands, at least 195.000 people suffer from forms of dementia, from which 60 to 70% is related to Alzheimer's disease (Alzheimer Nederland, 2006). At the age of 85 and higher, more than 30% are diagnosed with dementia (Alzheimer Nederland, 2006). The behavioural changes in Alzheimer patients can be difficult to cope with for the people that are close to them. According to Mace & Rabins (2001), emotional distress, frustration, anger and depression are some of the emotions family members of people with Alzheimer's disease show when they care for their relatives.

Knowledge can be a burden, in particular if the knowledge involves possible threats to well-being. Researchers have spotlighted a human tendency to avoid, ignore, and/or deny information, especially in the context of health care (Case et al., 2005). The appraisal of person-related health care information gives rise to coping mechanisms that try to reduce the possible threat, and this can give rise to strong emotional behaviours (see Leventhal, 1971). Avoiding information is closely linked to feelings of anxiety and fear, and if people pay attention to the information, it will cause mental discomfort (Case et al., 2005). This is related to the stress response described by Ursin and Eriksen (2004). When the affective value of an expected event is unattractive, e.g. running the risk of a certain health condition or disease, a conflict could appear with peoples' general 'stay healthy' set values. Because the Alzheimer domain is generally more personally related to seniors than the art domain, and because personal health care related information could cause discrepancies between set- and actual values, it is expected that seniors experience higher levels of stress when searching for Alzheimer information than for art information.

Most seniors that participated in the previous experiments had intermediate experience with PCs and the Internet. Seniors without experience were not able to perform the search tasks and had to be excluded from the sample. Because only about half of the people in the age group of 65-75 years have access to a PC and the Internet (CBS, 2006), a considerable number of the people in this age group will not have the possibility to become experienced with PCs and the Internet. A substantial number of the seniors that participated in the previous studies had difficulties operating a standard mouse and a QWERTY keyboard (see chapter 4, this thesis). Some of these seniors mentioned that the location of the characters on the QWERTY keyboard are odd, they preferred an alphabetical layout. Others used their index finger to click on the buttons on the screen. Such types of input uncertainty could easily be overcome. In the present experiment, system complexity is reduced for the inexperienced seniors by replacing the mouse with a touchscreen and the QWERTY keyboard with an ABCDE keyboard.

In chapter 4 (this thesis) it was found that the stress experience in senior computer interaction is dependent on induced expected task performance. The main goal of the research reported in this chapter is an extra validation (replication) of the findings reported in chapter 4, within the Alzheimer domain, and by means of text retrieval.

5.2 Experiment

5.2.1 Method

Participants

Thirty-two seniors were invited for participation in the experiment at Brunswijck Care Centre and surrounding senior apartments, in Eindhoven (The Netherlands). The experiment was conducted in the apartments of the seniors, to reduce external environmental effects on stress. In total, thirty seniors signed up for participation of which seven cancelled participation, and six seniors were excluded from the sample because of disruptions by relatives or low willingness to carry on with the procedure.

Seventeen seniors remained, 9 male and 8 female (mean age 80.24, $SD = 4.91$), and were paid €10,-. Criteria for participation were: familiarity with the English language, and having no or little experience with PCs. Ten of the participants had no experience with computers and the Internet from which four did not want to use a computer, one found himself too old, three found it too hard, and one just started working with a computer. Six had little experience with computers and the Internet (only text editing and Google), and one had intermediate experience with computers and the Internet (text editing, other programs, web site visits). Eight of the participants had an education level lower than High School.

Material

The experiment was conducted on a Dell latitude D800 laptop with a Pentium M processor and with a 15.4” wide screen (1280x800-pixel resolution). The low complexity system and the high complexity system (see figure 5.1 & 5.2) were used as the platforms on which the participants performed search tasks. The design and functionality of the two complexity systems was similar to the One-Box interface described in chapter 2 (this thesis).

The low complexity system (see figure 5.1) consisted of the laptop fitted with a mounted touchscreen and ABCDE keyboard, without a mouse. The ABCDE keyboard was a QWERTY keyboard on which key stickers were attached in ABCDE layout, from top left to bottom right. On the Enter key, a ‘ZOEK’ (search) sticker was placed, and on the Space bar, a ‘SPATIE’ (space) sticker was placed. All other keys were hidden by black stickers. The software used to change the assignment of the keys was Microsoft Keyboard Layout Creator. On screen, the ‘ZOEK’ (search) and ‘WISSEN’ (delete) buttons had a 1.27 cm (height) by 2.31 cm (width) surface area, the font was Arial in bold and 14-point size. The text fill-in box had a 0.59 cm (height) by 8.73 cm (width) surface area, and Arial font, 14-point size. The font of the Alzheimer textual information was MS Sans Serif, with a 13-point size, presented in a 17.02 cm (height) by 20.72 cm (width) surface area.

The high complexity system (see figure 5.2) consisted of the laptop with a QWERTY keyboard, a standard mouse, and a standard screen (no touchscreen). The size of the buttons and fonts was smaller than that of the low complexity system, and the text on the buttons was in English compared to Dutch on the low complexity system. The ‘Search’ and ‘Delete’ buttons had a 0.65 cm (height) by 1.94 cm (width) surface area, the font was MS Sans Serif, and 8-point size. The text fill-in box had a 0.54 cm (height) by 4.79 cm (width) surface area, and MS Sans Serif font, 8-point size. The font of the Alzheimer textual information was MS Sans Serif, with an 8-point size, presented in a 17.02 cm (height) by 20.72 cm (width) surface area.

Users could search for the information in the low- and high complexity systems by filling in one or more search words. The search words are compared and matched with the word by a document index that is present in both systems; stop words (e.g. the, and, it, etc.) are ignored. The content in both Alzheimer text retrieval systems consisted of the ‘question and answer’ information that is provided in Dutch on the ‘Alzheimer Nederland’ web site (Alzheimer Nederland, 2006)[7].

The tasks were written on paper. A stopwatch was used to measure task time in seconds. All instructions were read from pre made printed cards.

[7] The information is also available in folder format and can be ordered, without charge, on the web site.

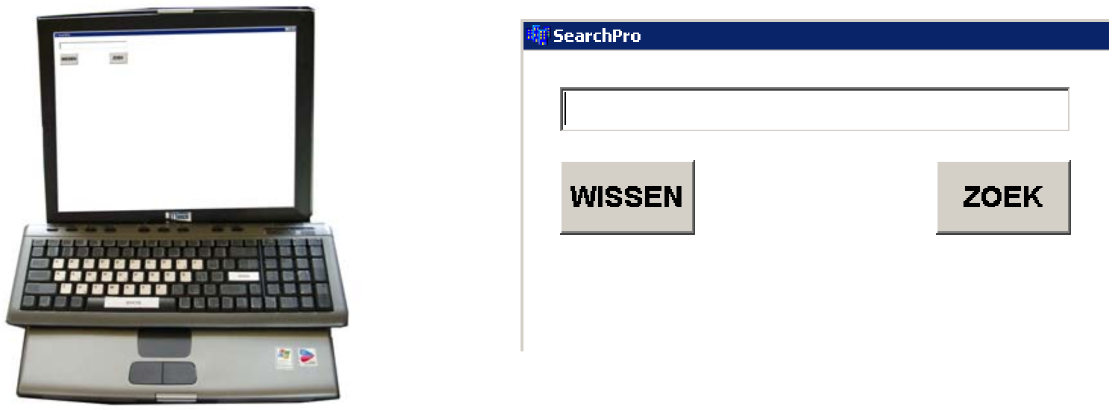

Figure 5.1: Low complexity system: laptop input configuration (left), and Alzheimer text retrieval system (right)

Figure 5.2: High complexity system: laptop input configuration (left), and Alzheimer text retrieval system (right)

Tasks

Two matching sets of two practice search tasks were to be completed on the two search interfaces. The practice search tasks served to present feedback- and feed-forward information about task performance. After that, two matching sets of ten search tasks were to be completed on both systems. The ten search tasks served to compare the task performance of the seniors on the low- and high complexity systems, and to examine the effect of induced expected task performance and complexity on stress and expected task performance over time. The matched search tasks in each set concerned information from the same text file. Half of the tasks in each task set concerned information that was located at the beginning of a text file, and half was located at the end of a text file. If the information was located at the beginning of a text file, then the matched search task concerned information that was located at the end of the same text file. An example of a search task: “With what type of test is the working of the nervous system examined?”

Dependent variables

Acute Stress

Subjective stress was measured by having participants fill in a rating scale on four attributes: gestressed (stressed), gespannen (tensed), geprikkeld (irritated), and zenuwachtig (nervous). These attributes had to be rated on an 11-point rating scale (0 = not at all, 10 = totally), see chapter 3 & 4.

Expected Task Performance

Before the practice tasks, (pre) expected task performance was measured by the question: "Now, you will receive four practice tasks that are executed with a computer program with which you can search for information about Alzheimer’s. How likely do you think you will perform the tasks correctly?" that had to be rated on an 11-point scale ranging from 0 (very unlikely) to a 100 (very likely) with 10 unit intervals (see chapter 4).

After the practice tasks, (post1) expected task performance was measured by the question: "Now, you will receive ten tasks that are executed on the computer program with which you just searched for information about Alzheimer’s. How likely do you think you will perform the tasks correctly?" that also had to be rated on the 11-point scale.

After the ten search tasks, (post2) expected task performance was measured by the question: "If you would work with this computer program in the future, how likely do you think you could perform the tasks correctly?" that also had to be rated on the 11-point scale.

Task performance

Task performance was used as a dependent variable to check if the low complexity system is more effective and efficient (ISO 9421-11) to retrieve information about Alzheimer than the high complexity system. Effectiveness was measured by the number of search tasks performed correctly, and efficiency was measured by total task time. When a participant was unable to perform a task within the time limit of the specific task, the time limit was coded as the efficiency score.

Manipulations

Complexity was induced by a low- and high complexity Alzheimer text retrieval system (see figure 5.1 and 5.2), and expected task performance was induced by fictitious negative and positive feedback, comparison, and feed-forward information. To induce low expected task performance, a performance graph was shown in which fictitious low performance of the participant was compared with fictitious high performance of other persons, and the participant was told by the experimenter that the performance was worse than the performance of the other persons. See chapter 4 for a precise description of the manipulation.

Design

A within-subjects design was used. The participants performed two matched sets of similar tasks on the two systems. The presentation order of the two task sets and the two systems was counterbalanced.

Procedure

The participants had received an informed consent form by letter. On arrival, the signed informed consent form was collected and then, the participants filled in a questionnaire to collect their characteristics. Next, information was given about the experimental procedure and it was explicitly stated that they were allowed to stop the experiment whenever they wanted. They had five minutes to relax, subjective stress was measured (pre-measurement), and their expected task performance was measured (pre-measurement). After this, the participants were given a short introduction about the search tasks they had to perform, and how they could perform the search tasks. For the low complexity system, information was given about the functionality of the touchscreen, ABCDE keyboard and the buttons on the screen. For the high complexity system, information was given about the functionality of the mouse, QWERTY keyboard and the buttons on the screen. Information was given about the best way to find an answer to the search task question. They were told that 'less is more' and that unique search words can better be used than more common search words (e.g. for the search task "How many people above 80 years old have dementia?", '80 years old' is more unique than 'dementia'). This was followed by the two practice search tasks; both tasks had a time limit of one minute. After the practice tasks, negative or positive feedback was given about their performance, their expected task performance was measured again (post1-measurement), and after that subjective stress was measured (post1-measurement). Participants carried out the ten search tasks, with a one-minute time limit. This was followed by measuring their subjective stress (post2-measurement). After this, expected task performance was measured (post2-measurement). The participants had a five-minute break to relax and after that, the procedure was repeated for the other system. See figure 5.3 for an overview of the procedure. In the debriefing, any possible emotional distress during the experiment was reduced by restorative communication. Information was given about the nature and goal of the research. The performance feedback screen was shown again and it was explicitly stated that the feedback given was fictitious. The total experiment took about one hour and forty minutes per participant.

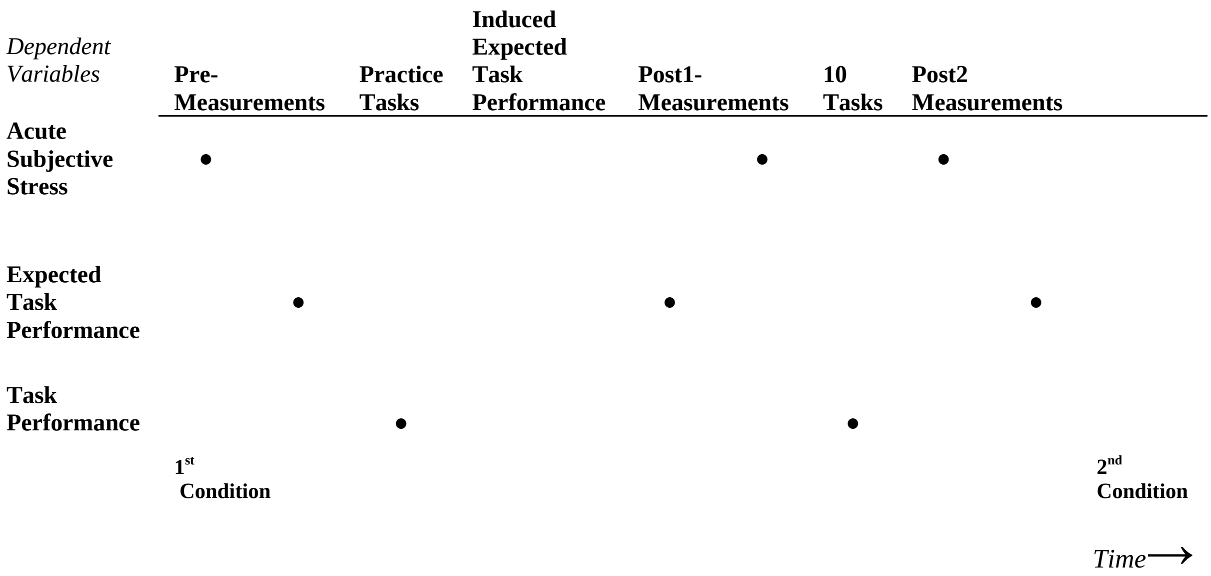

Figure 5.3: Procedure timeline

5.2.2 Results

Stress scale reliability

The mean Cronbach's alpha of the six repetitions of the subjective stress measurement scales was 0.89 (range 0.83-0.94); this is high enough to use the mean value of the four items as a subjective stress variable.

Task performance by task sets

Because two different, but matching, task sets were used, a paired samples t-test was performed to compare the two task sets by the number of tasks performed correctly and the total time it took to perform the tasks. If no difference is found, task set can be disregarded as a between-subjects factor in the following analyses. No significant difference was found between the task sets on the number of tasks performed correctly [$t(16) = .084$, $p = .934$], and no significant difference was found between the task sets on the total time it took to perform the tasks [$t(16) = .163$, $p = .873$]. See table 5.1 for means and standard deviations.

Table 5.1: Averages (*M*) and standard deviations (*SD*) of task performance on task set 1 and task set 2

		Task Set	
		1	**2**
Number of Tasks correct	*M*	3.83	3.77
(max.=10)	*SD*	2.43	2.80
Task Time	*M*	497.3	500.3
(in seconds, max.=600)	*SD*	62.5	80.0

The effect of temporal order on task performance, expected task performance and subjective stress

To examine a possible temporal effect on the dependent variables, a paired samples t-test was performed to compare the task performance of the preceding and following condition by the number of tasks performed correctly, and the total time it took to perform the tasks. (see table 5.2 for averages and standard deviations). No significant difference was found between the number of tasks performed correctly in the following condition compared to the preceding condition [$t(16) = .463$, $p = .649$], nor in the time it took to perform the tasks between both conditions [$t(16) = .201$, $p = .843$].

Table 5.2: Average task performance (*M*) and their standard deviations (*SD*) for the preceding and following condition

		Condition	
		preceding	**following**
Number of Tasks correct	*M*	3.59	3.88
(max.=10)	*SD*	2.76	2.39
Task Time	*M*	500.7	496.9
(in seconds, max.=600)	*SD*	79.1	63.6

A repeated-measures ANOVA was performed with time (3 levels: pre, post1 and post2) as a within-subject factor, with expected task performance and subjective stress as the dependent variables. Condition (preceding and following) was used as a between-subject factor.

No significant main effect was found of condition on expected task performance [$F(1, 32) = .162$, $p = .690$] and of condition on subjective stress [$F(1, 32) = .006$, $p = .940$]. No significant interaction effects were found between time and condition on expected task performance [$F(1, 32) = .452$, $p = .506$] and between time and condition on subjective stress [$F(1, 32) = .010$, $p = .921$]. See table 5.3 for averages and standard deviations.

Table 5.3: Averages (*M*) and standard deviations (*SD*) of expected task performance and subjective stress per condition (preceding and following) and time (pre, post1, and post2)

		Expected Task Performance			Subjective Stress		
Condition		**Pre**	**Post1**	**Post2**	**Pre**	**Post1**	**Post2**
Preceding	*M*	46.5	47.7	45.3	1.43	1.66	1.57
	SD	13.2	14.8	29.6	1.77	2.06	1.92
Following	*M*	45.9	49.4	50.6	1.82	1.54	1.44
	SD	14.2	20.2	25.1	1.97	1.97	2.12

5.2.3 Manipulation checks

Complexity

A paired samples t-test was performed to compare the task performance of the participants on the two practice tasks when complexity was low or high. No significant difference was found between the number of practice tasks performed correctly on the low- and high complexity systems [$t(16) = 2.073$, $p = .055$], and between the total time it took to perform the practice tasks [$t(16) = 1.770$, $p = .096$]. Table 5.4 shows means and standard deviations. The task performance difference on the practice tasks between the low- and high complexity systems were close to significant. However, the mean practice task performance on both complexities was low, so it is doubtful that the expected task performance manipulation was influenced favouring one of the complexities.

Table 5.4: Averages (*M*) and their standard deviations (*SD*) of task performance on the practice tasks on the low-and high complexity systems

		Complexity	
		Low	**High**
Number of Tasks correct	*M*	0.53	0.18
(max.=2)	*SD*	0.72	0.39
Task Time	*M*	109.4	116.9
(in seconds, max.=120)	*SD*	15.0	6.69

A paired samples t-test was performed to compare the complexity of the two interfaces that were used, by the number of tasks performed correctly, and the total time it took to perform the tasks (see table 5.5 for averages and standard deviations). Significantly more tasks correct were performed when complexity was low compared to high [$t(16) = 2.971$, $p = .009$]. The participants were significantly faster in performing the search tasks when complexity was low compared to high [$t(16) = 3.587$, $p = .002$].

Table 5.5: Averages (*M*) and standard deviations (*SD*) of task performance on the low- and high complexity systems

		Complexity	
		Low	**High**
Number of Tasks correct	*M*	4.59	2.88
(max.=10)	*SD*	2.09	2.74
Task Time	*M*	476.8	520.8
(in seconds, max.=600)	*SD*	59.3	76.0

Induced expected task performance

A repeated-measures ANOVA was performed with time (2 levels: pre and post1) as a within-subject factor, with expected task performance as the dependent variable. Induced expected task performance (positive and negative feedback) was used as a between-subject factor.

A significant interaction effect was found between time and feedback on expected task performance [$F(1, 32) = 10.5$, $p = .003$], with an effect size (partial η^2) of .247 and an observed power of .881. After the participants received negative feedback their expected task performance dropped, while after positive feedback there was a gain in expected task performance (see table 5.6 for averages and standard deviations).

Table 5.6: Averages (*M*) and standard deviations (*SD*) of induced expected task performance (positive and negative) by pre- and post1 expected task performance

Induced Expected Task Performance		Expected Task Performance: Pre	Post1
Positive	*M*	44.74	54.74
	SD	15.41	15.77
Negative	*M*	48.00	40.67
	SD	10.82	16.68

5.2.4 The effect of complexity and feedback on subjective stress and expected task performance

A repeated-measures ANOVA was performed with complexity (2 levels: high and low) and time (3 levels: pre, post1 and post2) as within-subject factors, with subjective stress and expected task performance as the dependent variables. Feedback group (positive-positive, negative-negative, positive-negative, negative-positive) was used as a between-subject factor.

No significant main effect was found of complexity on subjective stress [$F(1, 13) = 1.81$, $p = .202$], with an effect size (partial η^2) of .122 and an observed power of .238. Participants did not experience more stress when complexity was high than when complexity was low (see table 5.7 for averages and standard deviations).

Table 5.7: Averages (*M*) and standard deviations (*SD*) of subjective stress per level of complexity (Low and High) and time (pre, post1, and post2)

Complexity		Subjective Stress: Pre	Post1	Post2
Low	*M*	1.54	1.57	1.28
	SD	1.82	1.95	1.86
High	*M*	1.71	1.63	1.74
	SD	1.94	1.61	2.15

A significant main effect was found of feedback group on expected task performance [$F(3, 13) = 4.56$, $p = .022$], with an effect size (partial η^2) of .513 and an observed power of .766. Post-hoc tests with a Sidak correction indicated that the expected task performance of the participants in the pos-neg feedback group was significantly higher than the expected task performance of the participants in the neg-pos feedback group [$t(13) = 3.550$, $p = .021$].

A significant main effect was found of complexity on expected task performance [$F(1, 13) = 4.76$, $p = .048$], with an effect size (partial η^2) of .268 and an observed power of .524. Participants expected to perform better when complexity was low than when complexity was high (see table 5.8 for averages and standard deviations).

Table 5.8: Averages (*M*) and standard deviations (*SD*) of expected task performance per level of complexity (low and high) and time (pre, post1, and post2)

		Expected Task Performance		
Complexity		**Pre**	**Post1**	**Post2**
Low	*M*	44.71	50.00	54.12
	SD	14.63	17.68	20.63
High	*M*	47.65	47.06	41.77
	SD	12.51	17.59	31.87

An additional analysis was performed to examine the effect of induced expected task performance (positive and negative feedback) on subjective stress before and after the feedback was given.

A repeated-measures ANOVA was performed with time (2 levels: pre and post1) as a within-subject factor, with subjective stress as the dependent variable. Induced expected task performance (positive and negative feedback) was used as a between-subject factor.

A significant interaction effect was found between time and feedback on subjective stress [$F(1, 32) = 6.68$, $p = .015$], with an effect size (partial η^2) of .173 and an observed power of .708. Participants' subjective stress increased after they received negative feedback and decreased after they received positive feedback (see table 5.9 for averages and standard deviations).

Table 5.9: Averages (*M*) and standard deviations (*SD*) of subjective stress as a function of induced expected task performance (positive and negative) and time (pre and post1)

Induced Expected Task Performance		Subjective Stress	
		Pre	**Post1**
Positive	*M*	1.68	1.32
	SD	1.84	1.64
Negative	*M*	1.55	1.97
	SD	1.93	2.36

5.2.5 Correlations of post2 subjective stress and post2 expected task performance with task performance

Correlations were calculated to explore the relation between the total task times (efficiency) and the total tasks correct (effectiveness) of the tasks performed and the stress experience and task performance expectancies of seniors after the tasks were performed. It is expected that the correlations are similar to the correlations that were found in chapter 4. A positive linear relation is expected between post2 subjective stress and total task times, and between post2 expected task performance and total tasks correct. In addition, a negative linear relation is expected between post2 subjective stress and total tasks correct, and between post2 expected task performance and total task times.

Bivariate correlations indicated a significant Pearson correlation at the 0.05 level (2-tailed) between efficiency and post2 subjective stress [r = .369], and between effectiveness and post2 subjective stress [r = -.389]. In addition, a significant correlation was found at the 0.01 level between efficiency and post2 expected task performance [r = -.485], and between effectiveness and post2 expected task performance [r = .512]. Subjective stress increases with decreasing task performance, and expected task performance increases with increasing task performance.

5.2.6 T-test of mean difference in subjective stress between the Alzheimer and art domain

The mean subjective stress data of the research presented in chapter 4 (this thesis) were compared with the mean subjective stress data of the research presented here. The total mean of the pre-, post1- and post2-measurements are compared to examine if seniors experience higher levels of stress when they search for Alzheimer information compared to art images (see table 5.10 for means and standard deviations).

Table 5.10: Averages (*M*) and standard deviations (*SD*) of subjective stress per information domain (Alzheimer and art) and time (pre, post1, and post2)

		Subjective Stress			
Information Domain		**Pre**	**Post1**	**Post2**	**Total**
Alzheimer (*n* = 34)	*M*	1,63	1,60	1,51	1.58
	SD	1,86	1,99	1,99	1.87
Art (*n* = 60)	*M*	2,38	2,68	2,59	2.55
	SD	2,04	2,12	2,10	1.90

A t-test with unequal n's was performed to compare the total mean subjective stress of the Alzheimer and art domain. Seniors were significantly more stressed when the searched for art images than when they searched for Alzheimer information [$t(92) = 5.893$, $p < .000$].

5.2.7 Qualitative results

Alzheimer

All participants knew somebody close that was diagnosed with Alzheimer's or had dementia. The Alzheimer information elicited personal emotional reactions, e.g. one of the participants mentioned: "I really hope I will never become like this, how awful." The knowledge about Alzheimer's disease of the participants seemed intermediate to high. More than four of the participants had a health related profession in the past. These participants already gave an answer to a question before searching, and did not always agree with the information presented. Other participants also disagreed with some of the information that was available in the system. Five participants had lost their partner who had Alzheimer, and in contrast with the other participants, these participants fully agreed with the information that was available in the system. One of the participants had Alzheimer's disease (middle stage), without the experimenter's knowledge beforehand.

Stress

Most participants mentioned that they were not stressed during the experiment because they were in their home environment. Three of the participants were stressed during the whole experiment, they said that they were always stressed in situations in which they were tested and had to perform. One participant mentioned that she was stressed during a computer course she followed, and even used Valium

(Diazepam) to relax and to be able to sleep. The participant who was diagnosed with Alzheimer's became irritated at one certain moment during the experiment when he had to find an answer to the question: "What type of character changes can appear?" He first denied that character changes appear, and found the question quite odd. Afterwards, he told, and his wife confirmed, that he showed aggressive behaviour among people who did not understand his condition, which resulted in social isolation.

Complexity

It is interesting to note that the participants with no computer experience were immediately able to interact with the low complexity system. From their comments it appeared that they were pleased with the alphabetical layout of the keyboard. It should be mentioned that the touchscreen was not always used when participants had to press the search button; the Enter key search button was also used. The font size of the high complexity system was hard to read for some of the participants. It resulted in typing errors (misspellings), and the participants moved closer to the screen. Five participants had difficulties operating the mouse, and many found it hard to find the right keys on the QWERTY keyboard. About all seniors who had no computer experience mentioned that they also had no typewriter experience. Participants who had experience with QWERTY keyboards found it difficult to operate the ABCDE keyboard, although they were still able to perform search tasks within the time limit. One of the participants had bad eyesight in combination with tunnel vision. Tunnel vision resulted in difficulties locating the mouse pointer. If the mouse was moved, he could not follow the pointer in the direction it was moved. He was pleased with the touchscreen because he could use his own finger as an interaction device.

5.3 Discussion

Again, as in chapters 3 and 4, it has been found that complexity has no effect on subjective stress in senior computer interaction. When seniors interact with a high complexity system, their subjective stress is not higher than when they interact with a low complexity system. Induced expected task performance has a significant effect on subjective stress. Subjective stress increases when seniors receive negative feedback- and feed-forward information about their performance. Expected task performance could be induced by performance feedback. When seniors receive negative feedback about their task performance, expected task performance decreases, whereas when they receive positive feedback, expected task performance increases. Complexity did have an effect on expected task performance. Seniors' expected task performance was higher when they performed search tasks on the low complexity system compared to the high complexity system. Seniors were significantly more efficient and effective in performing search tasks on the low complexity system than on the high complexity system. Although a significant difference was found, the mean difference in efficiency and effectiveness between both complexities was not large. In addition, the mean number of search tasks performed correctly on the low complexity system was lower than half of the total number of search tasks presented. It is likely that more tasks would have been performed correctly if the participants had had more time to perform the tasks. The ultimate goal is a task performance score close to a 100% correct. Accessible systems contribute to this goal, but training programs to improve the computer skills of seniors are also necessary to enhance task performance (see Mead et al., 1999).

The seniors without computer- and typewriter experience were pleased with the alphabetical layout of the keyboard. The extra iterative design step provided seniors, with no computer experience and typing skills, the ability to successfully interact with a digital information retrieval system. To increase accessibility, it appeared to be helpful to provide large fonts (14-point size) and buttons, add a touchscreen, and replace the QWERTY keyboard by an ABCDE keyboard. Not only was the QWERTY layout replaced by an ABCDE layout, all other keys (e.g. F1, Backspace, etc.), except for a ‘Search’ and ‘Space’ key, were hidden by black stickers, which likely decreased distraction. The large fonts and buttons in the design of the low complexity system very likely helped to overcome age-related vision problems. Because the touchscreen was not always used to perform a task, and because touchscreens cannot be purchased at regular computer shops and are expensive, keyboard overlays could be sufficient for the inexperienced senior computer user. Keyboard overlays have been used in music production and photo editing design for many years now and can be effective and efficient to access deep menu structures and shortcut key combinations. Keyboard overlays that provide the user with the necessary functions to perform standard tasks could replace the mouse as an input device. For inexperienced senior computer users, a low cost keyboard overlay for example for an operating system desktop, a document editor, or an Internet search engine, could overcome accessibility problems. It should be noted that the Alzheimer text retrieval system needed little text input (only a couple of search words) from the user to retrieve information. For text editing, a QWERTY layout may be preferred, because alphabetical layouts have been found to be inferior compared to the QWERTY layout (see Kroemer, 2001), even for the novice user. Noteworthy, the last QWERTY vs. ABCDE keyboard layout studies were conducted in the beginning of the 1980s (see Kroemer, 2001), at a time command-line interfaces were far more common than GUI interfaces. Aside from text editing, little textual input is needed to operate current GUIs. Seniors without any experience with QWERTY keyboards could be given the opportunity to work with ABCDE keyboards or other native language specific alphabetical layouts from start to overcome a possible psychological barrier, or computer anxiety, to learning and using computers.

In contrast with the experiment presented in chapter 4 (this thesis), no order effects were found. The experience that was gathered to perform search tasks on one of the systems in the preceding condition did not seem to contribute to a better task performance on the other system in the following condition. This could have been caused by the larger difference in the input devices of the low- and high complexity systems, e.g. knowledge gathered about the location of the characters on the QWERTY keyboard cannot be transferred to locating characters on the ABCDE keyboard. It is likely that significant interaction effects would have been found between time and condition on expected task performance and subjective stress if more seniors had participated in the experiment. From the descriptive statistics, it appeared that expected task performance increases and subjective stress decreases in the second (following) experimental condition, while in the first (preceding) experimental condition expected task performance stays the same over time and there is a small increase in subjective stress.

Task performance was found to be related to subjective stress and expected task performance. Similar to the findings of chapter 4 (this thesis) it was found that subjective stress increases with decreasing task performance. In addition, a linear relation was found between efficiency and effectiveness, and expected task performance, which was also found in chapter 4. Expected task performance increases with task performance. In senior computer interaction, high subjective stress and low expected task performance is paired with low efficiency and effectiveness.

It was not expected that the stress experience of seniors would be lower when performing search tasks on an Alzheimer text retrieval system compared to an art image retrieval system. However, the subjective stress of the seniors was already lower before they searched for Alzheimer information compared to art images. A possible explanation is that the seniors experienced higher levels of stress during the experiment presented in chapter 4 (this thesis), because blood pressure (BP) was measured immediately before subjective stress was measured. BP measurements could increase stress experiences, as it is intrusive and related to health problems and diseases. The seniors that participated in both experiments differed in age and computer experience. The mean age of the seniors that participated in the experiment presented in chapter 4 was lower and the level of computer experience was higher than the seniors that participated in the experiment presented in this chapter. It is unclear if the difference in age and computer experience between both groups influenced the difference in the stress they experienced. The significant difference of subjective stress in seniors between both domains could be explained by difference in the environments in which the experiments were conducted. The experiment presented in chapter 4 was conducted in a shared living room, while the experiment presented here was conducted in the homes of the seniors. In general, people feel more comfortable in their own homes, with their own belongings around them. This is related to primary appraisal (see Folkman et al., 1986), because in most cases there is less harm to self-esteem when people fail to perform tasks in their home environment compared to a less familiar external environment. In addition, if there would emerge a situation in the home environment that is perceived as harmful to the resident, then it is quite easy to prevent any additional harm by removing the stressful stimulus out of the home environment, which is related to secondary appraisal.

The Alzheimer domain did seem to cause seniors to be personally and emotionally involved during the experiment. The domain elicited narratives about friends and relatives, some deceased, who had been diagnosed with Alzheimer's disease. A considerable number of seniors disagreed with the Alzheimer information that was presented after a search task. This could have been caused by the minimal presentation of the Alzheimer information, e.g. no references to studies, authors, and official Alzheimer organizations. Perceived credibility has been found to influence an individuals' level of trust in health web sites (see Briggs, Burford, De Angeli & Lynch, 2002). By only presenting textual information, without references, a lack of trust in the information could easily have taken place. On the other hand, the seniors that had lost their partners with Alzheimer's disease fully agreed with the information presented. It is possible that it is not (only) a lack of trust in the information, but an exposure of mental discomfort and denial. The reason that the senior who was diagnosed with Alzheimer disagreed with a particular question could have been caused by one of the general problems of the behaviour of Alzheimer's patients, which is the denial of problems (Gordeau & Hiller, 2005).

Stress clearly plays a role in senior computer interaction. When seniors receive negative feedback about their performance, subjective stress increases. The complexity of a digital information system does not affect subjective stress in senior computer interaction. In addition, the information domain of a retrieval system does not seem to influence stress experiences, although the Alzheimer domain does seem to cause seniors to be personally and emotionally related to the information presented. Because seniors mentioned that they were not stressed during the experiment, it is likely that seniors experience little stress during computer interaction within the home environment, but of course, without negative performance feedback.

Acknowledgements

Thanks to the personnel and inhabitants of Brunswijck and surrounding senior apartments for their hospitality.

Chapter 6: General discussion

The research presented in this thesis focused on increasing accessibility and understanding stress, in senior computer interaction. The research is part of the NWO NARRATOR project, with the goal to gather functional and non-functional requirements for an information retrieval system that provides breast cancer patients with narratives of other patients and ways to communicate with each other. To gather these requirements, we started out within the art domain with non-patients. This was done to prevent unnecessarily hindering patients with disease-related information provided by a retrieval system that still contains design flaws (see also Blanson Henkemans et al., 2007). The experiments have shown, by iterative design, how to provide high accessibility for seniors. Given that a large group of breast cancer patients are at the age of 60 or more (RIVM, 2007), the developed retrieval systems can serve as an accessible platform for health related information retrieval. In addition, the experiments shed light on why seniors are stressed during computer interaction. Positive performance feedback information can reduce stress and increase expected task performance. This knowledge might be valuable in the design of health related information retrieval systems to provide stress-vulnerable patients a low stress computer interaction experience.

In the following subchapter, an overview of the main findings and requirements for NARRATOR are presented first, before a more general overview is given and implications for practice and future research are discussed.

6.1 Overview of results

Seniors have been found to face difficulties learning and using new media (see Charness & Holley, 2004; Docampo Rama, 2001; Freudenthal, 2001; Mead et al., 1999; Mead et al., 2000; Morrell et al., 2002; Xie, 2002). To overcome problems with accessibility, digital information systems should be engineered with reference to the knowledge and capabilities of seniors. The experimental evaluations presented in chapter 2 show how to design an accessible retrieval system for seniors. The first experiment demonstrated that young experienced users benefited from a cognitively engineered image retrieval system, while the seniors with less experience were not. The seniors showed high levels of stress, and eventually the experiment had to be stopped. Cognitive engineering is an iterative process, and for the senior computer users extra steps were necessary to provide high accessibility. The error analysis that was performed shed light on the problems the seniors faced during interaction. As a result of the error classifications, three different redesigns were constructed, intended to overcome task comprehension-, location- and system comprehension errors. All three redesigns provided seniors with high accessibility and low stress. The last experiment presented in chapter 2 showed that certain interface functions were flawed on the first developed prototype. With these functions, all age groups performed worse on the first prototype than on the redesigned prototype. In conclusion, the developed category based prototype provided high accessibility for all age groups to seniors. Although this is a valuable improvement, it remained unclear why the seniors were stressed during the first experimental evaluation.

Chapter 3 focused on the effect of complexity on stress in senior computer interaction. The results showed that when seniors performed search tasks on a high complexity system, their acute physiological- and subjective stress was not higher than when they performed search tasks on a low complexity system. The heart rate and subjective stress of the seniors dropped significantly after search tasks were performed, independent of the complexity of both systems. Because no significant task performance difference was found as a function of the system complexity, the accessibility of the high complexity system and low complexity system were similar. A probable cause for this similarity is the way the search tasks were presented on the screen. Seniors had to search for art images, and the tasks (e.g. 'Find a painting of Rembrandt') were presented in a panel on the screen. Most seniors were confused about the task presentation and perceived it as part of the interface. In addition, the seniors browsed through the images instead of using the category based search options. Automated task presentations should be avoided in senior computer interaction research.

Chapter 4 focused on the effect of complexity and expected task performance on stress in senior computer interaction. A theoretical model of stress in senior computer interaction was developed. The model consists of two factors, complexity and expected task performance, which could influence the stress experience and response in senior computer interaction. It was expected that when seniors interact with a high complexity system that their acute subjective- and physiological stress is higher and expected task performance is lower, than when they interact with a low complexity system. In addition, when seniors' expected task performance is manipulated by negative performance feedback information, it was expected that acute subjective and physiological stress would be higher than when they receive positive performance feedback information. It was indeed found that when seniors receive negative performance feedback, expected task performance decreases and subjective stress increases. When seniors interact with a high complexity system, their expected task performance declined while when they interact with a low complexity system expected task performance increased. Seniors were significantly more efficient and effective in performing search tasks on the low complexity system than on the high complexity system. Stress plays a role in senior computer interaction when performance feedback is taken into account. Although it was found that the complexity of a digital information system does not play a role in stress in senior computer interaction (see figure 6.1 for the effect sizes that were found), actual task performance seemed to be related to subjective stress. Subjective stress increased with decreasing task performance.

Chapter 5 focused on stress in senior computer interaction within the health domain. An Alzheimer's text retrieval system was engineered, and an extra iterative design step was made to increase accessibility for the inexperienced seniors. From the previous experiments it appeared that seniors with no or little experience with PCs had accessibility problems with the redesigned art image retrieval system. These seniors were confused about the QWERTY layout of the keyboard, and some seniors used their index fingers to click the button on the screen rather than on the mouse. To overcome this type of input uncertainty, and to lower complexity, the QWERTY keyboard was replaced by an ABCDE keyboard, and a touchscreen was added. The goal was an extra validation of the experiment reported in chapter 4, plus the influence of personal emotional related information on acute stress in senior computer interaction. Again, complexity had no effect on subjective stress in senior computer interaction (see figure 6.1), but subjective stress increased with decreasing task performance. Furthermore, it was found that when seniors receive negative feedback about their performance, expected task performance decreased and subjective stress increased. Although the Alzheimer's domain elicited personal emotional reactions in seniors, it did not produce higher stress experiences in seniors than when they searched for art images. The additional iterative design step provided seniors, with no to little PC experience with high accessibility. An interesting finding was that many seniors did not trust the Alzheimer information that was presented, which could have been caused by the minimal presentation of the information. Seniors that had lost a partner with Alzheimer's disease fully agreed with the information presented, so it is possible that is not (only) a lack of trust in the information presented, but an exposure of mental discomfort and denial.

Stress plays a role in senior computer interaction. Acute subjective stress increases when seniors receive negative performance feedback information. The complexity of a system has no effect on subjective stress (see figure 6.1), but it does have an effect on the expectancies seniors have about their future task performance and the efficiency and effectiveness with which tasks are performed. By iterative design and by taking the experience and abilities of seniors into account, most seniors, even seniors without any computer experience, are able to interact successfully with a digital information system.

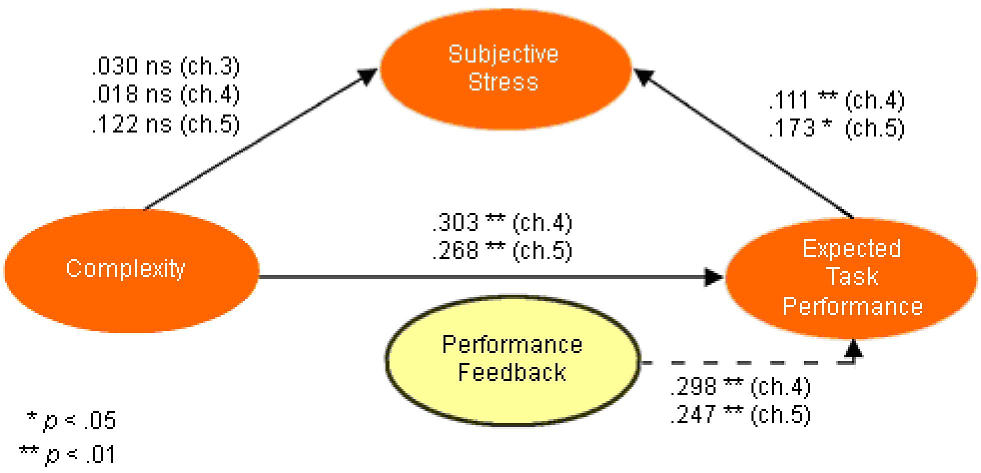

Figure 6.1: Theoretical model of acute stress in senior computer interaction with effect sizes (partial η^2) found in chapter 3, 4, and 5. Expected task performance was induced by performance feedback. The effect sizes presented between expected task performance and subjective stress represent the effect of the performance feedback manipulation on subjective stress.

The experiments presented in this thesis contributed to understanding one of the unexplored topics within senior computer interaction. According to Hamborg and Grief (2003) there is a lack of evidence about the direct impact of software design on stress, and Hawthorn (2000) suggested an exploration of how stress affects, and is affected by, interface use for older users. The empirical evidence presented here shows that the complexity of an interface does not affect stress in senior computer interaction. Feedback and feed-forward task performance information do influence subjective stress and the expectancies seniors have about their own performance. This is in line with Ursin and Eriksen (2004), who state that stress depends on the individual appraisal of the situation, which is based on previous experience and expectations of the outcome. Expectations of the outcome can be induced by performance feedback and feed-forward information that influences stress in senior computer interaction. When seniors receive negative performance feedback, subjective stress increases and expected task performance decreases, while positive performance feedback decreases subjective stress and increases expected task performance. Before the seniors received feedback- and feed-forward information about their performance, only a fourth of the practice tasks were performed correctly on the low- and high complexity systems. Though seniors' absolute performance was low, they still expected to perform well after positive feedback was given. If expectancies of future performance drop between two successive measurements, then there likely is a conflict between a set value and an actual value, which triggers a stress response. This is in accordance with the Cognitive Activation Theory of Stress (Ursin & Eriksen, 2004). Pre expected task performance that was measured in the experiments presented in chapter 4 & 5 (this thesis) is similar to a set value. Post expected task performance is similar to an actual value. A relation has been found between task performance and expected task performance, so absolute task performance could contribute to the actual value, but it appears that perceived expectancies induced by performance feedback have a considerable impact on the actual value. Slegers (2006) suggested that psycho education could reduce computer anxiety in inexperienced seniors. Positive feedback and feed-forward information could serve as an effective type of psycho education to reduce computer anxiety in inexperienced seniors. The importance of positive feedback in senior computer interaction has been suggested by Czaja et al. (2006). According to them, to overcome computer anxiety, it is critical to ensure that older people receive encouraging feedback during training and experience some level of success.

It was not expected that complexity had no influence on stress in senior computer interaction. From a HCI or Cognitive Engineering perspective, it seems plausible that the stress experience of users is higher when they interact with a high complexity system than with a low complexity system, because a high complexity system offers the user less control and means to cope with task complexity demands. To quote Formosa (2005): "It happens all too often. Poorly designed products and interfaces cause anger and stress, or worse, are misunderstood and misused when people are under stress." From the experiments presented in this thesis it appears that the stress experience of seniors is not higher when they perform search tasks on a poorly designed product (high complexity) than on a better designed product (low complexity). It was found that the complexity of a product does not influence stress, but the outcome expectancies or the expected task performance, induced by feedback information about performance does influence stress. Similar results have been found by Eriksen and Ursin (1999) on subjective health complaints in a study of 1060 people working at the Norwegian postal service. They found that the objective possibility of having control, which was operationalized by, among other things, skills and task variety, had a lower impact on subjective health complaints than coping. Coping was defined as an expectancy of positive outcome and was operationalized by, among other things, being optimistic about outcomes and approaching difficult situations.

If complexity has no effect on stress in senior computer interaction, why was stress observed in the first experiment presented in chapter 1 (this thesis)? The seniors could have experienced stress because of an interplay between low computer experience, negative expectancies about task performance, low task performance, and the environment in which the experiment was conducted. The participants had little to intermediate computer experience. When people have little knowledge about how to interact with a computer, then they have few resources available to deal with task complexity demands, which could have an influence on stress. A more likely cause, in accordance with the findings in this thesis, is that the seniors with little computer experience had negative expectancies about their task performance, which could increase stress (see Ursin & Eriksen, 2004). In addition, task performance was almost zero. In chapter 4 and 5 (this thesis) it was found that there is a linear relation between task performance and subjective stress. Subjective stress increases with decreasing task performance. When task performance is low then it is possible that seniors experience higher subjective stress. Similar to the effect of negative induced task performance on subjective stress, low perceived task performance could have an influence on stress. Because task performance was low, self-esteem might have been at stake, which could result in a stress experience, but also because the experiment was conducted at a University lab and not in the home environment of the participants. Self-esteem is related to primary appraisal (see Folkman et al, 1986), and has been found to have a moderating influence on psychological and physiological stress (Rector & Roger, 1997). In general, people feel more confident in their home environment compared to a lab environment. In a lab environment, it is more difficult to withdraw from a possible harmful encounter compared to a home environment. This is in line with the research of Ursin and Eriksen (2004) who state that when the perceived probability of avoiding an unpleasant encounter approaches zero, then people experience a state of helplessness, which increases stress. The seniors who participated in the experiment could have experienced a state of helplessness. They could also have experienced a state of hopelessness, because although the participants were able to interact to some extent with the art image retrieval systems, i.e. there was some control, about all actions had a negative outcome. Hopelessness arises when there is a very high probability that available responses bring results of high negative affective value (Ursin & Eriksen, 2004).

Although complexity did not influence stress in senior computer interaction directly, it did have an effect on expected task performance and the efficiency and effectiveness with which tasks can be performed. Expectancies about future performance increased when seniors interacted with a low complexity system, while expectancies decreased when they interacted with a high complexity system. In other words, their self-efficacy increased when complexity was low. By iterative design steps, high accessibility was provided for seniors, even for inexperienced seniors. The design strategy presented in this thesis requires early involvement of seniors, experimental evaluations in which the design is compared with existing exemplars and iterative design by using error analysis and thorough user observations that provide requirements for future designs. It is shown that it is important to strive for minimal interaction, with large fonts (14-point size) and buttons with a clear function in the design. For the inexperienced senior, accessibility can be increased by an ABCDE keyboard layout and by making the mouse superfluous by means of a touchscreen and system specific keyboard keys (e.g. a ‘Search’ key sticker on the Enter key). The inexperienced seniors were pleased with the ABCDE layout, which reopens the discussion about QWERTY vs. ABCDE and other layouts (see Kroemer, 2001). The knowledge that was gathered from the experiments is presented in table 6.1 as a guideline for the design of accessible systems for seniors. Some of the guidelines are not unique in the sense that they can be traced back to other guidelines and interface requirements for senior computer users (e.g. Czaja & Lee, 2003; Czaja &

Lee, 2007; Fisk et al., 2004; Mead et al., 2000; Morrell et al., 2002; Welford, 1985). The requirements[8] are in conformance with the work that has been carried out already, but add ideas that need further exploration.

Table 6.1: Interface guidelines for senior computer interaction

Design

Guideline	*Improves/Overcomes*
Use large fonts (14-point size)	Ease of Reading
Use large buttons (1.3 * 2.3 cm surface area)	Navigation
Use native language in the design	Function Comprehension
Use 1 information layer	Distraction, Disorientation
Provide smallest amount of functions necessary for task	Distraction
Provide positive feedback information	Stress Experience, Expected Task Performance
Consider the use of an ABCDE keyboard layout	Keyboard Layout Comprehension
Replace mouse by touchscreen or by short keys on keyboard or keyboard overlays	Navigation, Efficiency
Avoid scrolling (also in drop-down menus)	Navigation, Disorientation
Avoid Up Down buttons	Navigation

6.2 Requirements for NARRATOR

The knowledge that was gathered in this thesis can be used to provide interface requirements for NARRATOR; a narrative retrieval system for breast cancer patients, and other health related retrieval systems. The interface requirements consist of most of the guidelines that are presented in table 6.1. Because a large group of breast cancer patients are seniors (see RIVM, 2007) it is important to provide large fonts and buttons in the design. Because Dutch seniors can face difficulties when information is presented in English, it is required that the functions and content in the design is presented in Dutch. Avoid using more than one information layer and provide the smallest amount of functions necessary for a retrieval task. In addition, for successful navigation, avoid scrolling and 'Up Down' buttons. A category based interface, or one-box interface (see figures 2.3 and 2.4, chapter 2) could serve as a platform to provide narratives of patients. Overberg et al. (2007) found that patients are mainly interested in fellow patients' experiences about how to cope with emotions, the impact of cancer in daily life, and physical discomforts. Similar topics were found in a preliminary study to gather metadata within the breast cancer domain. Four diaries of patients and the Viva forum (VIVA, 2003) were used to count the most frequent used words by patients, which resulted in 86 words that could be divided under six categories: 'ziekte' (disease), 'activiteiten' (activities), 'ziekenhuis' (hospital), 'lichaam' (body), 'naasten' (fellow (wo)men),

[8] Most of the requirements are based on heuristics and should rather not be treated as hard facts.

and ‘gevoelens’ (feelings). Figure 6.2 shows a prototype of a possible NARRATOR search interface, in which the above requirements are implemented, with examples of the concepts that could be used in the design (in Dutch). It should be mentioned that the prototype is an example of a possible NARRATOR search interface and that the interface still needs to be experimentally evaluated with the intended user group.

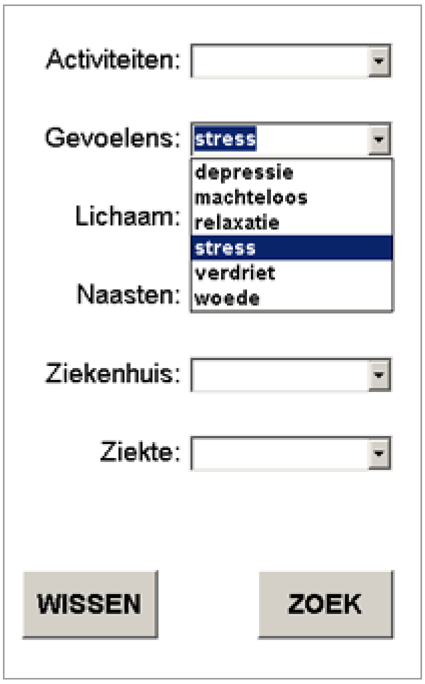

Figure 6.2: Example of a NARRATOR search interface

In addition to the above requirements it is important that a retrieval system that provides narratives of breast cancer patients is designed in such a way that is does not bring about additional stress and negative feelings. In this thesis it was found that subjective stress in senior computer interaction can be reduced by positive performance feedback information, and that is important to avoid negative performance feedback information. Future studies could focus on how positive feedback information could be implemented in health related retrieval systems (see 6.3.2 for an example). For a breast cancer narrative retrieval system, and other health information retrieval systems, it is advisable to present feedback messages in a positive sense, or at least avoid negative error messages (e.g. ‘wrong input’).

Trust is another important factor to consider in health related information retrieval. In chapter 5 (this thesis) it was found that many seniors did not trust the information that was presented by the Alzheimer retrieval system. Although, the seniors who had lost their partner who had Alzheimer’s disease fully agreed with the information presented. It is questionable why many seniors had a lack of trust in the information. They might have disagreed with the information because they recognized some of the dementia and Alzheimer problems from their own behaviour, which resulted in mental discomfort and denial, or the presentation of the information lacked in credibility. The Alzheimer information was presented in plain text (black on white), with a 13-point size, without any references to Alzheimer organizations. In online transactions it has been found that young adults' trust in music e-tailers is influenced by the reputation of the vendor (Metzger, 2006). Within the health domain, it has been found that the design of a health web site and the perceived credibility influences an individuals' level of trust (Briggs et al., 2002). The perceived credibility of the Alzheimer retrieval system could have been low, caused by the 'non-professional' minimalistic way the health related content was presented, and the lack of references to trustworthy health organizations, which ultimately could have decreased the seniors' level of trust in the information. References to high-reputation health organizations or hospitals could increase the level of trust in health related digital information. For a narrative health related retrieval system like NARRATOR, references to health organizations could be of help to increase trust. But, the perceived authenticity of a patients' narrative might need other cues to increase the trust other patients have in the narrative, e.g. the addition of a profile with information of the name, age, etc. of patients. Future studies could focus on the importance of trust cues in narrative health-related retrieval systems, and the type of cues that can increase trust, and how these cues should be implemented in the design.

6.3 Future research and implications for practice

From the research that was executed, implications for practice and future research can be given. These are divided into future stress research in senior computer interaction, and future feedback research and implications for feedback design in senior computer interaction.

6.3.1 Future stress research in senior computer interaction

In the following paragraphs, future stress research is subdivided into stress measurements, the information domain, length of stress reduction, influence of other humans, and interface complexity.

Stress Measurements: No acute physiological stress responses were found in the experiments presented in this thesis, while acute subjective stress experiences were clearly found. Instead of measuring blood pressure (BP) and heart rate (HR), in future research, other physiological measurements (e.g. level of Cortisol, Heart Rate Variability, palmar sweating) could be explored to study stress in senior computer interaction, or BP could be measured continuously by means of the TNO Portapress (when the device is validated). But, as mentioned by Doublet (2000), many physiological measures of stress are not necessarily indicative of the presence or existence of stress. Perceived or subjective stress measures might well be the most relevant indicators of stress, as they have been found to be more reliable than some of the physiological measures (Shostak & Peterson, 1990). Because most physiological stress measurement methods are intrusive (e.g. BP measurement), future studies in senior computer interaction could focus on other unobtrusive means to measure stress, aside from subjective measurements, to provide reliable results and to safeguard the well-being of subjects. For example, stress could be measured unobtrusively by automated emotion recognition systems that detect emotional speech and facial expressions, although much research is necessary to validate these types of systems (see Truong, Van Leeuwen & Neerincx, 2007). In this thesis, significant interactions were found between feedback and subjective stress, but the mean subjective stress values were low. A probable cause for these low values could be the relatively safe home environment in which the experiments were executed. Furthermore, the stress manipulation by induced expected task performance was quite mild for ethical reasons, and in general it is difficult and improper to create real emergency situations that threaten an organism's physiological well-being (see Bourne & Yaroush, 2003). Future research could explore, for example questionnaires or diary studies, the level of stress and frustration seniors experience, and when they experience stress during computer interaction in practice. In the introduction of this thesis, it was mentioned that the use of the concept of stress in daily life is relatively new in the Netherlands. Diary studies could elicit what concepts seniors use to express a stress experience, because the use of the concept of stress and related negative stress concepts might be less common and/or improper among seniors than among young adults.

Information Domain: The Alzheimer’s domain did not cause seniors to experience higher levels of stress compared to the art domain. However, it is likely that people experience higher levels of stress when searching for health information that concerns their own health. Future research could try to explore what the level of impact is on stress when seniors search for information that considers their own health, and how positive feedback information could support these particular users. It is important to consider clinical ethics when doing this type of research, to safeguard the well-being of patients.

Length of Stress Reduction: As proposed by Czaja et al. (2006) and in line with the research presented in this thesis, future research could focus on the effect of positive feedback in senior computer interaction on stress reductions in the long run, and how these successful manipulations of attitudinal variables generalize to new developments in technology. Additional subjective stress measurements and expected task performance measurements, e.g. weeks, months and even years after the experiment, could provide knowledge whether the effects of positive feedback are maintained over time. In addition to these measurements, it would also be interesting to measure motivation. Are seniors more motivated to explore computers and interact with them after positive feedback? Furthermore, did positive feedback change their views about computers, i.e. did it lower computer anxiety? In the experiments presented in this thesis acute stress was measured. Acute stress is not harmful to people’s health except if it is sustained over many years. It would be interesting to conduct a longitudinal study in which seniors' stress experiences with current technologies is studied to gather knowledge about a possible chronic nature of stress experiences with these technologies that could certainly have an effect on seniors' health.

Influence of Other People: It is likely that a part of the stress seniors experience and the fear seniors have of modern technologies is (un)intentionally caused by other people (e.g. friends and relatives). Female seniors have been found to report more computer anxiety than males (see Czaja et al., 2006; Karavidas et al., 2005), which could be traced back to classic male-female roles in which males take care of technology. Future research could focus on the influence of friends, relatives and colleagues on stress in senior computer interaction and seniors' anxiety and fear of computer technologies, and how technology could contribute to overcoming this anxiety and fear.

Complexity: The difference in actual task performance of seniors on the two system complexities was significant in most experiments but not always large. Although it is unlikely to find an effect of complexity on stress in senior computer interaction (this thesis), it would be interesting to execute an experiment in which one of the complexity systems is highly inaccessible for all seniors. The task performance of seniors on this system should be close to zero, similar to the first experiment presented in chapter 2 (this thesis). When all actions have negative results, people experience a state of hopelessness, which increases stress (see Ursin & Eriksen, 2004). It should be mentioned that it could be difficult to gather stress data in such an experiment because the participants might stop the experiment before any stress data are gathered. Trimmel, Meixner-Pendleton and Haring (2003) found that longer system response times when searching for information on the Internet caused a stress response in young adults. This stress response was not related to participants' expertise. Uncertainty grows and HR increases when system response times increase. The response times of the systems that were used in the experiments in this thesis were negligible. Future research could compare a system with long response times, e.g. 30 seconds, to a system with 5-second response times to study if similar effects will be found for seniors as for the young adults in the study of Trimmel et al. (2003). It should be mentioned that a likely cause for users to experience stress when response times are longer is due to outcome expectancy. For example, people who are used to a broadband Internet connection will likely get frustrated and stressed when they have to use narrowband (e.g. a 56k modem) to download files or to surf the Internet. Given that seniors are less experienced (see Brynin et al., 2004) and less confident about their judgements within the computer domain (Marquié et al., 2002) than young adults, it is possible that differences in complexity of current computer technologies do not have an influence on stress but differences in technologies of different generations. Docampo Rama (2001) found that the number of errors people make using complex user interfaces and the strategy they use are technology generation specific, in particular when confronted with a new product for the first time. A difference between old- and current technologies might result in a difference in a stress experience. If a senior has a lot of experience with old technology, e.g. an analogue alarm clock, and has to use current technology, e.g. a digital alarm clock, then the performance expectancies in setting a digital alarm clock might be negative from the start. The digital alarm clock might in fact be a low complexity system for all, but the anxiety and negative expectancies to use current technologies could result in seniors experiencing stress when confronted with current technologies for the first time. Positive feedback and feed-forward performance information could overcome this anxiety.

6.3.2 Future feedback research and implications for practice

"If you can design the interface of a system or a product in such a way that the user feels in control of the user situation, the product can actually reduce stress!" (Nilsen, 2005).

In this thesis, examples were given of seniors who have low expectancies about successfully learning about and using computers. Some seniors who followed an Internet course excused themselves from their exam with the given reason 'exam fear' (Krijnen, 2004), and a senior that participated in the experiment presented in chapter 4 even used Valium during a computer course to be able to relax and sleep. Positive feedback systems and virtual teachers could support these seniors in overcoming their fears. In addition, these systems could also be beneficial for the less able seniors and the seniors who live in remote areas, i.e. those seniors who cannot be physically present at computer courses.

The way the positive and negative performance feedback information was given in the experiments presented in this thesis was artificial compared to regular computer usage, and the feedback information was false to serve as an experimental manipulation. The challenge in future research would be to implement positive feedback information in digital information systems in such a way that the user takes the information into consideration, that the information is effective in increasing expected task performance and decreasing stress, and that the feedback information could adapt to the level of experience and knowledge of the user.

It would not be too difficult to design feedback information in such a way that the user considers the information. To ensure visibility and accessibility for seniors, the presentation of the feedback information should follow the guidelines of table 6.1, e.g. large fonts and native language. The work of Nass, Moon, Fogg, Reeves and Dryer (1995) could be used to endow personality types (e.g. submissive and dominant) in the feedback information that are recognizable by users. Nass et al. (1995) found that submissive users prefer a submissive computer, while dominant users prefer a dominant computer. It is possible to emulate personality in computers, and feedback information could be altered in such a way that the messages match the personality type of the user. Jussim et al. (1995) found that people high in self-esteem accept positive feedback more readily than negative feedback and people low in self-esteem accept negative feedback more readily than positive feedback. Future research could explore how seniors with different levels of self-esteem accept negative- or positive feedback more readily. Knowledge from the field of Persuasive Technologies (see IJsselsteijn et al., 2006) and agent-based interactive instructive systems (Masthoff, 1997), could be used to design effective feedback that positively changes the expectancies seniors have about their own performance. A virtual coach or teacher could be implemented that provides seniors with positive feedback and feed-forward messages. Two virtual teachers with different personalities (submissive and dominant) could be presented to the user first from which he or she can select the one preferred. The virtual teacher could present performance outcome and expectancy graphs that are similar to the graphs used in chapter 4 & 5 (this thesis) to increase expected task performance and to reduce subjective stress. The work of Masthoff (1997) could further be used to design situated agents that provide feedback that adapts to the individual knowledge and level of experience of the user. For seniors, it is preferred that this feedback information is positive. Masthoff (1997) used a level of enthusiasm and disappointment in a feedback agent, and even found that users showed a tendency to make mistakes on purpose to get the virtual teacher enthusiastic when answering correctly afterwards. A virtual teacher could be designed for the specific tasks that seniors want or need to perform on a computer, e.g. e-mail and online banking. Inexperienced senior computer users could be given more basic functionality support by the virtual teacher, while experienced senior computer users could be given advanced functionality support, and after task mastery, the feedback could be presented less often or might even be absent in future tasks. It is important that the feedback can be voluntarily turned on and off by the user, because it might become annoying (e.g. Microsoft Paperclip), and might consequently even have a harmful effect on satisfactory computer interaction.

6.4 Conclusion

The results from the experiments presented in this thesis show that the complexity of a digital system does not have a direct effect on acute stress in seniors. Seniors' acute subjective stress is not higher when they perform search tasks on a high complexity system compared to a low complexity system. However, stress plays a role in senior computer interaction. Subjective stress increases and expected task performance decreases when seniors receive negative performance feedback and feed-forward information. Complexity does have an effect on the efficiency and effectiveness seniors perform tasks and on the expectancies seniors have about their task performance. Within the art- and health domain, similar results have been found. The experimental evaluations have shown how to design an accessible retrieval system for seniors, even for the inexperienced senior computer user. With the knowledge gathered from the experiments, system- and interface engineers could implement aiding that induces positive expected task performance, in such a way that seniors perceive high expected task performance and experience low stress.

It is promising news that for the upcoming six years, the European Union will spend one hundred and fifty million Euros in ICT for seniors, to support independent living, longer-lasting health and workability (Automatisering Gids, 2007). Although this is a very positive development, it is important that todays and future seniors will always have the choice in learning and using new information and communication technologies. Following this, I would like to conclude with some words from Lazarus & Lazarus (2006), "Children as well as elderly need to believe that they are valuable as persons, deserve respect from others as individuals, and should also respect themselves without being fooled about what they are able and unable to do and be." In senior computer interaction, a small contribution can be made to fulfil these conditions by the design of accessible systems and the addition of positive performance feedback, to contribute to seniors' ability to participate in today's digital area.

References

AllRefer. *Neurasthenia, Psychology And Psychiatry*. Retrieved April 18, 2007, from http://reference.allrefer.com/encyclopedia/N/neurasth.html

Alzheimer Nederland. Retrieved December 3, 2006, from http://www.alzheimer-nederland.nl/

Anderson, J.D., & Pérez-Carballo, J. (2001). The nature of indexing: how humans and machines analyze messages and texts for retrieval. Part I: Research, and the nature of human indexing. *Information Processing and Management, 37*, 231-254.

Andrade, E.N. da C. (1960). Robert Hooke, F.R.S. (1635-1703). *Notes and Records of the Royal Society of London (1938-1996), 15*.

Atkinson, B.M.C. (2006). Captology: A Critical Review. In IJsselsteijn, De Kort, Midden, Eggen, & Van den Hoven (Eds.) *First International Conference on Persuasive Technology for Human Well-Being* (pp. 171-182). Berlin: Springer.

Automatisering Gids. *Brussel wil ouderen langer op de been houden met ICT*. Retrieved August 15, 2007, from http://automatiseringgids.sdu.nl/ag/nieuws/nieuws/toon_nieuwsbericht.jsp?di=344854

Baddeley, A. (1997). *Human Memory – Theory and Practice* (Revised Edition). East Sussex: Psychology Press Ltd.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioural change. *Psychological Review, 84*, 191-215.

Belkin, N.J. (2002). Helping People Find What They Don't Know. *Communications of the ACM, 43*, 58-61.

Berger, D.M. (1973). The Return of Neurasthenia. *Comprehensive Psychiatry, 14*, 557-562.

Birdi, K.S., & Zapf, D. (1997). Age differences in reactions to errors in computer-based work. *Behaviour & Information Technology*, *16*, 309-319.

Blanson Henkemans, O.A., Rogers, W.A., Fisk, A.D., Neerincx, M.A., Lindenberg, J., & Van der Mast, C.A.P.G. (2007). *An Adaptive Computer Assistant for the Supervision of Older Diabetics' Self-Car*e. Manuscript submitted for publication.

Boon, L. (1996). *Geschiedenis van de psychologie* (6th ed.). Meppel: Boom.

Bouma, H., & Graafmans, J. (Eds.) (1992). *Gerontechnology.* Amsterdam: IOS Press.

Bourne, L.E., & Yaroush, R.A. (2003). Stress and cognition: A cognitive psychological perspective (NASA Grant Number NAG2-1561). Retrieved July 20, 2007, from http://human-factors.arc.nasa.gov/eas/download/non_EAS/Stress_and_Cognition.pdf

Bouwhuis, D.G. (1991). Dialogue Constraints in Instruction. In R.J. Beun, M. Baker & M. Reiner (Eds.), *Dialogue and Instruction – Modeling Interaction in Intelligent Tutoring Systems*, 142, (pp. 15-30). NATO ASI Series F: Computer and Systems Sciences.

Bouwhuis, D.G. (1992). Aging, perceptual and cognitive functioning and interactive equipment. In: H.Bouma & J.A.M. Graafmans (Eds.), *Gerontechnology*, (pp. 93-112). Amsterdam: IOS Press.

Bouwhuis, D.G. (2000). Parts of life: configuring equipment to individual lifestyle. *Ergonomics, 43*, 908-919.

Bouwhuis, D.G. (2003). Design for Person-Environment Interaction in Older Age: a Gerontechnological Perspective. *Gerontechnology, 2,* 232-246.

Briggs, P., Burford, B., De Angeli, A., & Lynch, P. (2002). Trust in online advice. *Social Science Computer Review, 20*, 321-332.

Bronswijk, J.E.M.H. (2006). Gerontechnology motivation. *Gerontechnology, 5*, 65-67.

Brynin, M., Raban, Y., & Soffer, T. (2004). *e-Living: Life in a Digital Europe*, Chapter 5: The New ICTs: Age, Gender and the Family. Retrieved July 30, 2007, from http://www.eurescom.de/e-living/deliverables/e-liv-D14-Ch5-Family.pdf

Callister, R., Suwarno, N.O., & Seals, D.R. (1992). Sympathetic activity is infuenced by task difficulty and stress perception during mental challenge in humans. *Journal of Physiology, 454*, 373-387.

Cannon, W.B. (1932). *The Wisdom of the Body*. New York: W.W. Norton & Company Inc.

Case, D.O., Andrews, J.E., Johnson, J.D., & Allard, S.L. (2005). Avoiding versus seeking: the relationship of information seeking to avoidance, blunting, coping, dissonance, and related concepts. *Journal of the Medical Library Association, 93,* 353-362.

CBS. (2006). *Onderzoek ICT gebruik bij personen*. Retrieved November 20, 2006, from http://statline.cbs.nl/StatWeb/table.asp?PA=71098ned&D1=a&D2=0&D3=(l-11)-l&DM=SLNL&LA=nl&TT=2

Charness, N., Holley, P. (2004). The new media and older adults: Usable and useful. *American Behavioral Scientist, 48,* 416-433.

Clifford, W., & Scott, M. (1956). Sigmund Freud, 1856-1939. *Canadian Medical Association Journal, 74*, 744-745.

CNET. *An aplhabetical keyboard.* Retrieved October 23, 2006, from http://news.com.com/2300-1041-6005621.html

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2^{nd} ed.). Hillsdale, NJ: Erlbaum.

Cooper, C.L., & Dewe, P. (2004). *Stress – A Brief History*. Oxford: Blackwell Publishing.

Cremers, A.H.M., & Neerincx, M.A. (2004). Personalisation Meets Accessibility: Towards the Design of Individual User Interfaces for All. In: *User-Centered Interaction Paradigms for Universal Access in the Information Society. Lecture Notes in Computer Science* (pp.119-124). Berlin: Springer.

Czaja, S.J., & Sharit, J. (1993). Stress Reactions to Computer-Interactive Tasks as a function of Task Structure and Individual Differences. *International Journal of Human-Computer Interaction, 5*, 1-22.

Czaja, S.J., & Lee, C.C. (2003). Designing Computer Systems for Older Adults. In J.A. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook – Fundamentals, Evolving Technologies and Emerging Applications* (pp. 425). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Czaja, S.J., Charness, N., Fisk, A.D., Hertzog, C., Nair, S.N., Rogers, W.A., & Sharit, J. (2006). Factors Predicting the Use of Technology: Findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and Aging, 21*, 333-352.

Czaja, S.J., & Lee, C.C. (2007). The impact of aging on access to technology. *Universal Access in the Information Society, 5*, 341-349.

dableducational. *Blood Pressure Monitors-Validations, Papers and Reviews*. Retrieved May 2, 2007, from http://www.dableducational.org/sphygmomanometers.html#non_invasive

Deakin, J., Aitken, M., Robbins, T., & Sahakian, B.J. (2004). Risk taking during decision-making in normal volunteers changes with age. *Journal of the International Neuropsychological Society, 10,* 590-598.

Degner, L.F., & Sloan, J.A. (1992). Decision making during serious illness: What role do patients really want to play? *Journal of Clinical Epidemiology, 45*, 941-950.

De Greef, H.P., & Neerincx, M.A. (1995). Cognitive Support: Designing Aiding to Supplement Human Knowledge. *International Journal for Human-Computer Studies*, *42,* 531-571.

Diamond, D.M., Campbell, A.M., Park, C.R., Halonen, J., & Zoladz, P.R. (2007). The Temporal Dynamics Model of Emotional Memory Processing: A Synthesis on the Neurobiological Basics of Stress-Induced Amnesia, Flashbulb and Traumatic Memories, and the Yerkes-Dodson Law. *Neural Plasticicty, 2007*, 1-33.

Dickinson, A., & Dewsbury, G. (2006). Designing computer technologies with older people. *Gerontechnology, 5,* 1-3.

Docampo Rama, M. (2001). *Technology generations handling complex user interfaces.* Doctoral dissertation. The Netherlands: Technische Universiteit Eindhoven.

Dohms, J.E., & Metz, A. (1991). Stress – mechanisms of immunosuppression. *Veterinary Immunology and Immunopathology, 30*, 89-109.

Doublet, S. (2000). *The Stress Myth*. Chesterfield: Science & Humanities Press.

Echt, K.V. (2002). Designing Web-Based Health Information for Older Adults: Visual Considerations and Design Directives. In Morrell, R.W. (Ed.) *Older Adults, Health Information and the World Wide Web* (pp. 73-74). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Ellis, R.D., & Allaire, J.C. (1999). Modeling computer interest in older adults: the role of age, education, computer knowledge, and computer anxiety. *Human Factors, 41*, 345-355.

Ellis, R.D. (2005). Ergonomics, human factors engineering, and Gerontechnology: From 'accommodation' to robust design. *Gerontechnology, 4*, 61-62.

Embrey, D.E. (1986). SHERPA: a systematic human error reduction and prediction approach. *International Topical Meeting on Advances in Human Factors in Nuclear Power Systems.* Knoxville: Tennessee.

Eriksen, H.R., Murison, R., Pensgaard, A.M., & Ursin, H. (2005). Cognitive activation theory of stress (CATS): From fish brains to the Olympics. *Psychoneuroendocrinology, 30*, 933-938.

Eriksen, H.R., & Ursin, H. (1999). Subjective health complaints: is coping more important than control?. *Work & Stress, 13*, 238-252.

Eyck, A., Geerlings, K., Karimova, D., Meerbeek, B., Wang, L., IJsselsteijn, W., De Kort, Y, Roersma, M., & Westerink, J. (2006). Effect of a Virtual Coach on Athletes' Motivation. In IJsselsteijn, De Kort, Midden, Eggen, & Van den Hoven (Eds.) *First International Conference on Persuasive Technology for Human Well-Being* (pp. 158-161). Berlin: Springer.

Fisk A.D., Rogers A.R., Charness N, Czaja S.J., & Sharit J. (2004). *Designing for Older Adults – Principles and Creative Human Factors Approaches.* Boca Raton: CRC Press.

Fitts, P.M. (1951). *Human engineering for an effective air navigation and traffic control system*. Washington, DC: National Research Council.

Folkman, S., Lazarus, R.S., Dunkel-Schetter, C., DeLongis, A., & Gruen, R.J. (1986). *Journal of Personality and Social Psychology, 50,* 992-1003.

Foltz, P.W. (1991). Models of Human Memory and Computer Information Retrieval: Similar Approaches to Similar Problems. *ICS Tech Report 91-03*.

Formosa, D. (2005). Design for Stress. *Design Management Review, 16*, 50-55.

Fox, S. (2001). *Wired Seniors – A fervent few, inspired by family ties*. Retrieved May 18, 2007 from http://www.pewinternet.org/pdfs/PIP_Wired_Seniors_Report.pdf

Fox, S. (2004). *Older Americans and the Internet*. Retrieved December 1, 2006, from http://www.pewinternet.org/pdfs/PIP_Seniors_Online_2004.pdf

Freudenthal, D. (2001). Age differences in the performance of information retrieval tasks. *Behaviour & Information Technology, 20,* 9-22.

Geerts, G., Heestermans, H., & Kruyskamp, C. (Eds.). (1984). *Van Dale's Groot woordenboek der Nederlandsche taal* (11th ed.). Utrecht/Antwerpen: Van Dale Lexicografie.

Gerhardt-Powels, J. (1996). Cognitive engineering principles for enhancing human-computer performance. *International Journal of Human-Computer Interaction, 8*, 189-211.

Google (2007). *"Information Retrieval"*. Retrieved September 27, 2007, from http://www.google.nl/search?hl=nl&q=%22information+retrieval%22&meta=

Gordeau, B.C., & Hiller, J. (2005). *Alzheimer's Essentials – Practical Skills for Caregivers.* Florida: Carma Publishing LLC.

Gulliksen, J., & Harker, S. (2004). The software accessibility of human-computer interfaces – ISO Technical Specification 16071. *Universal Access in the Information Society, 3*, 6-16.

Hamborg, K.C., & Grief, S. (2003). New Technologies and Stress. In M.J. Schabracq, J.A.M. Winnubst & C.L. Cooper (Eds.), *The Handbook of Work and Health Psychology (2nd ed.)* (pp. 211). Chichester: John Wiley & Sons Ltd.

Harrington, T.L., & Harrington, M.K. (2000). *Gerontechnology – Why and How.* Maastricht: Shaker.

Hawkley, L.C., & Cacioppo, J.T. (2004). Stress and the aging immune system. *Brain, Behavior, and Immunity, 18*, 114-119.

Hawthorn, D. (2000). Possible implications of aging for interface designers. *Interacting with Computers*, *12*, 507-528.

Hawthorn, D. (2003). How Universal is Good Design for Older Users? *Proceedings of the 2003 conference on Universal usability*, 38-45.

Hawthorn, D. (2006). Enhancing the contributions of older people to interface design. *Gerontechnology*, *5*, 4-15.

Hjortskov, N., Rissén, D., Blangsted, A.K., Fallentin, N., Lundberg, U., & Søgaard K. (2004). The effect of mental stress on heart rate variability and blood pressure during computer work. *European Journal of Applied Physiology*, *92*, 84-89.

Hoenkamp, E. (2003). Unitary Operators on the Document Space. *Journal of the American Society of Information Science, 54*, 314-320.

Hollnagel, E. (1998). *CREAM: Cognitive Reliability Error Analysis Method*. London: Elsevier.

Hollnagel, E., & Woods, D.D. (1983). Cognitive Systems Engineering: New Wine in New Bottles. *International Journal of Man-Machine Studies, 18,* 583-600.

Holmes, T.H., & Rahe, R.H. (1967). The Social Readjustment Rating Scale. *Journal of Psychosomatic Research, 11*, 213-218.

IJsselsteijn, W., De Kort, Y., Midden, C., Eggen, B., & Van den Hoven, E. (2006). Persuasive Technology for Human Well-Being: Setting the Scene. In IJsselsteijn, De Kort, Midden, Eggen, & Van den Hoven (Eds.) *First International Conference on Persuasive Technology for Human Well-Being* (pp. 1-5). Berlin: Springer.

Ingwersen, P. (1992). *Information Retrieval Interaction.* London: Taylor Graham.

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

Johnson, P. (1992). *Human-Computer Interaction: Psychology, Task Analysis and Software Engineering.* London: McGraw-Hill.

Judge, M. (Writer/Director). (1999). Office Space [Motion Picture]. USA: 20th Century Fox, Cubicle Inc.

Jussim, L., Yen, H., & Aiello, J.R. (1995). Self-Consistency, Self-Enhancement, and Accuracy in Reactions to Feedback. *Journal of Experimental Social Psychology, 31*, 322-356.

Karavidas, M., Lim, N.K., & Katsikas, S.L. (2005). The effects of computers on older adult users. *Computers in human behavior, 21*, 697-711.

Karel, M.J. (1997). Aging and Depression: Vulnerability and Stress Across Adulthood. *Clinical Psychology Review, 17*, 847-879.

Kelley, C., & Charness, N. (1995). Issues in training older adults to use computers. *Behaviour & Information Technology, 14*, 107-120.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* New York: Cambridge University Press.

Krijnen. (2004). *Jong helpt Oud Internet op.* Retrieved December 6, 2006, from http://krijnen.com/archives/000246.shtml

Kroemer, K.H.E., Kroemer, H.J., & Kroemer-Elbert, K.E. (1997). *Engineering Physiology – Bases of Human Factors/Ergonomics.* New York: Van Nostrand Reinhold.

Kroemer, K.H.E. (2001). Keyboards and keying – An annotated bibliography of the literature from 1878 to 1999. *Universal Access in the Information Society, 1*, 99-160.

Kruyskamp, C. (Ed.). (1976). *Van Dale's Groot woordenboek der Nederlandsche taal* (10^{th} ed.). 's-Gravenhage: Martinus Nijhoff.

Landauer, T.K., Foltz, P.W., & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes, 25*, 259-284.

Lazarus, R.S., & DeLongis, A. (1983). Psychological Stress and Coping in Aging. *American Psychologist, 38*, 245-254.

Lazarus, R.S., & Lazarus, B.N. (2006). *Coping with Aging*. New York: Oxford University Press.

Leventhal, H. (1971). Fear Appeals and Persuasion: The Differentation of a Motivational Construct. *American Journal of Public Health, 61,* 1208-1224.

Levine, E.G., Eckhardt, J., & Targ, E. (2005). Change in Post-traumatic Stress Symptoms Following Psychosocial Treatment for Breast Cancer. *Psycho-Oncology*, *14,* 618-635.

Lewis, J.R. (1995). IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. *International Journal of Human-Computer Interaction*, *7*, 57-78.

Lovallo, W.R., & Thomas, T.L. (2000). Stress Hormones in Psychophysiological Research – Emotional, Behavioral, and Cognitive Implications. In Cacioppo, J.T., Tassinary, L.G., & Berntson, G.G. (Eds.). *Handbook of Psychophysiology* (2^{nd} ed.) (pp. 344). Cambridge: Cambridge University Press.

Lundberg, U., & Frankenhaeuser, M. (1980). Pituitary-adrenal and sympathetic-adrenal correlates of distress and effort. *Journal of Psychosomatic Research*, *24*, 125-130.

Mace, N.L., & Rabins, P.V. (2001). *The 36-Hour Day – A Family Guide to Caring for Persons with Alzheimer Disease, Related Dementing Illness, and Memory Loss in Later Life.* New York: Warner Books Inc.

Making Your Web Site Senior Friendly. Retrieved March 1, 2007, from http://www.nlm.nih.gov/pubs/checklist.pdf

Marcellini, F., Mollenkopf, H., Spazzafumo, L., & Ruoppila, I. (2000). Acceptance and use of technological solutions by the elderly in the outdoor environment: findings from a European survey. *Zeitschrift für Gerontologie und Geriatrie, 33*, 169-177.

Marquié, J.C., Jourdan-Boddaert, L., & Huet, N. (2002). Do older adults underestimate their actual computer knowledge?. *Behaviour & Information Technology, 21*, 273-280.

Masthoff, J.F.M. (1997). An Agent-Based Interactive Instruction System. Doctoral dissertation. The Netherlands: Technische Universiteit Eindhoven.

Matthews, G., Campbell, S.E., Falconer, S., Joyner, L.A., Huggins, J., Gilliland, K., Grier, R., & Warm, J.S. (2002). Fundamental Dimensions of Subjective State in Performance Settings: Task Engagement, Distress, and Worry. *Emotion, 2*, 315-340.

Mayhorn, C.B., Lanzolla, V.R., Wogalter, M.S., & Watson, A.M. (2005). Personal Digital Assistants (PDAs) as Medication Reminding Tools: Exploring Age Differences in Usability. *Gerontechnology, 4,* 128-140.

McAlister, F.A., & Straus, S.E. (2001). Evidence based treatment of hypertension – Measurement of blood pressure: and evidence based review. *BMJ, 322*, 908-911.

McEwen, B., & Norton Lasley, E. (2002). *The end of stress as we know it.* Washington: Joseph Henry Press.

Mead, S.E., Sit, R.A., Rogers, W.A., Jamieson, B.A., & Rousseau, G.K. (2000). Influences of general computer experience and age on library database search performance. *Behaviour & Information Technology*, *19*, 107-123.

Mead, S.E., Batsakes P., Fisk, A.D., & Mykityshyn, A. (1999). Application of Cognitive Theory to Training and Design Solutions for Age-related Computer Use. *International Journal of Behavioral Development*, *23*, 553-573.

Mead, S.E., Lamson, N., & Rogers, W.A. (2002). Human Factors Guidelines for Web Site Usability: Health-Oriented Web Sites for Older Adults. In Morrell, R.W. (Ed.) *Older Adults, Health Information and the World Wide Web* (pp. 92). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Melenhorst, A.S. (2002). Adopting communication technology in later life: The decisive role of benefits. Doctoral dissertation. The Netherlands: Technische Universiteit Eindhoven.

Metzger, M.J. (2006). Effects of Site, Vendor, and Consumer Characteristics on Web Site Trust and Disclosure. *Communication Research, 33*, 155-179.

Morrell, R.W., Mayhorn, C.B., & Bennett, J. (2000). A Survey of World Wide Web Use in Middle-Aged and Older Adults. *Human Factors, 42*, 175-182.

Morrell, R.W. (Ed.). (2002). *Older Adults, Health Information and the World Wide Web.* Mahwah, New Jersey: Lawrence Erlbaum Associates.

Nap, H.H., De Greef, H.P., & Bouwhuis, D.G. (2005). Access for all by cognitive engineering. (CD-ROM). *Proceedings of the 5th International Conference of the International Society on Gerontechnology.* Nagoya, Japan. Nap, H.H., De Greef, H.P. & Bouwhuis, D.G. Access for all by cognitive engineering. (Abstract). *Gerontechnology*, *3*, 259.

Nardi, B.A. (Ed.). (1995). *Context and Consciousness: Activity Theory and Human-Computer Interaction.* Cambridge: Massachusetts Institute of Technology.

Nass, C., Moon, Y, Fogg, B.J., Reeves, B., & Dryer, D.C. (1995). Can computer personalities be human personalities? *International Journal of Human-Computer Studies, 43*, 223-239.

NGA. *National Gallery of Art*. Retrieved December 14, 2004, from http://www.nga.gov/search/index.shtm#artist

Neerincx, M.A., & De Greef, H.P. (1998). Cognitive Support: Extending Human Knowledge and Processing Capacities. *Human-Computer Interaction*, *13*, 73-106.

Neerincx, M.A. & Lindenberg, J. (in press). Situated cognitive engineering for complex task environments. In Schraagen, J.M. (Ed.) *Natural Decision Making & Macrocognition*. Ashley.

Neerincx, M.A., Lindenberg, J., Rypkema, J.A. & Van Besouw, N.J.P. (2000). A practical cognitive theory of Web-navigation: Explaining age-related performance differences. *CHI2000*.

Newell, A.F. (2006). Older People as a focus for Inclusive Design. *Gerontechnology, 4,* 190-199.

Nielsen, J.N. (1994). *Usability Engineering*. London: Academic Press.

Nielsen, J.N. (2002). *Usability for Senior Citizens*. Retrieved November 20, 2006, from http://www.useit.com/alertbox/20020428.html

Nilsen, K.L. (2005). Designing for stress - The need for guidelines when developing products or systems for use in stressful situations. Retrieved October 9, 2007, from http://design.ntnu.no/forskning/artikler/2005/artikkel_Kjetil_Loenne_Nilsen.pdf

Norman, D.A. (1982). Steps toward a cognitive engineering: Design rules based on analyses of human error. *Proceedings of the 1982 conference on Human factors in computing systems,* 378-382.

O'Brien, E., Asmar, R., Beilin, L., Imai, Y., Mallion, J.M., Mancia, G., Mengden, T., Myers, M., Padfield, P., Palatini, P., Parati, G., Pickering, T., Redon, J., Staessen, J., Stergiou, G., & Verdecchia, P. (2003). European Society of Hypertension recommendations for convential, ambulatory and home blood pressure measurement. *Journal of Hypertension, 21*, 821-848.

O'Brien, E., Waeber, B., Parati, G., Staessen, J., & Meyers, M. (2001). Blood pressure measuring devices: recommendations of the European Society of Hypertension. *BMJ, 22*, 531-536.

Overberg, R., Toussaint, P., & Zwetsloot-Schonk, B. (2006). Illness stories on the Internet – Features of websites disclosing breast cancer patients' illness stories in the Dutch language. *Patient Education and Counseling, 61*, 435-442.

Overberg, R., Alpay, L.L., Verhoef, J., & Zwetsloot-Schonk, J.H.M. (2007, February 2). Illness stories on the internet: what do breast cancer patients want at the end of treatment. *Psycho-Oncology*. Retrieved June 12, 2007, from http://www3.interscience.wiley.com/cgi-bin/fulltext/114109906/PDFSTART

Palmer, S.C., Kagee, A., Coyne, J.C., & DeMichele, A. (2004). Experience of Trauma, Distress, and Posttraumatic Stress Disorder Among Breast Cancer Patients. *Psychosomatic Medicine, 66,* 258-264.

Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction. An Approach to Cognitive Engineering.* New York: Elsevier Science Publishing Co.

Reason, J.T. (1987). Generic error modelling system: a cognitive framework for locating common human error forms. In J. Rasmussen, K.D. Duncan & J. Leplat (Eds.) *New Technology and Human Error*. Chichester: Wiley.

Rector, N.A., & Roger, D. (1997). The Stress Buffering Effects of Self-Esteem. *Personality and Individual Differences, 23*, 799-808.

Rettig, M. (1994). Prototyping for Tiny Fingers. *Communications of the ACM, April, 37.*

Rissén, D., Melin, B., Sandsjö, L., Dohns, I., & Lundberg, U. (2000). Surface EMG and psychophysiological stress reactions in women during repetitive work. *European Journal of Applied Physiology*, *2*, 215-222.

RIVM (2007). *Nationaal Kompas Volksgezondheid*. Retrieved April 19, 2007, from http://www.rivm.nl/vtv/object_document/o1492n17276.html

Robin, E.D., & Bromberg, P.A. (1959). Claude Bernard's Milieu Interieur Extended: Intracellular Acid-Base Relationships. *The American Journal of Medicine, 27*, 689-692.

Ropeik, D. (2004). The consequences of fear. *EMBO reports, 5*, 56-60.

Salas, E., Driskell, J.E., & Hughes, S. (1996). The Study of Stress and Human Performance. In Driskell & Salas (Eds.) *Stress and Human Performance* (pp. 1-46). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Sapolsky, R.M. (1996). Why stress is bad for your brain. *Science, 273,* 749-750.

Saracevic, T. (1996). Modeling interaction in information retrieval (IR): A review and proposal. *Proceedings of the American Society for Information Science, 33*, 3-9.

Sayers, H. (2004). Desktop virtual environments: a study of navigation and age. *Interacting with Computers, 16*, 939-956.

Sefelin, R., Tscheligi, M., & Giller V. (2003). Paper Prototyping – What is it good for? A Comparison of Paper- and Computer-based Low-fidelity Prototyping. *CHI 2003, New Horizons.*

Selye, H. (1970). The Evolution of the Stress Concept – Stress and Cardiovascular Disease. *The American Journal of Cardiology, 26*, 289-299.

SeniorWeb. *Organisatie*. Retrieved November 20, 2006, from http://www.seniorweb.nl/oversw/content.aspx?id=674&menuid=30

Sharit, J., Czaja, S.J., Perdomo, D., & Lee, C.C. (2004). A cost-benefit analysis methodology for assessing product adoption by older user populations. *Applied Ergonomics, 35*, 81-92.

Shneiderman, B. (1997). *Designing the User Interface – Strategies for Effective Human-Computer Interaction (3rd ed.).* Amsterdam: Addison-Wesley.

Shostak, B. B., & Peterson, R. A. (1990) . Effects of Anxiety Sensitivity on Emotional Response to a Stress Task. *Behaviour Research and Therapy, 28*, 513-521.

Simon, J.A. (2000). Does “white coat hypertension” increase the risk for any adverse outcome from hyptertension?. *Western Journal of Medicine, 172*, 47-48.

simPC. *Doorbraak in retail simPC.* Retrieved October 23, 2006, from http://www.simpc.com/node/93

Slegers, K. (2006). Successful Cognitive Aging: The use of computers and the Internet to support autonomy in later life. Doctoral dissertation. The Netherlands: Maatricht University.

Sterling, P. (2004). Principles of Allostasis: Optimal Design, Predictive Regulation, Pathophysiology, and Rational Therapeutics. In Schulkin, J. (Ed.) *Allostasis, Homeostasis, and the Costs of Physiological Adaptation* (pp. 17-64). Cambridge: Cambridge University Press.

Swain, M.J., & Ballard, D.H. (1991). Color Indexing. *International Journal of Computer Vision, 7,* 11-32.

Teigen, K.H. (1994). Yerkes-Dodson: A Law for all Seasons. *Theory & Psychology, 4*, 525-547.

TVM. *TigerTail Virtual Museum*. Retrieved December 14, 2004, from http://www.tigtail.org/TIG/S_View/TVM/X3/search.html#search

Topouchian, J.A., El Assaad, Mohamed, A., Orobinskaia, Ludmila, V., El Feghali, Ramzi, N., Asmar, & Roland, G. (2006). Validation of two automatic devices for self-measurement of blood pressure according to the International Protocol of the European Society of Hypertension: the Omron M6 (HEM-7001-E) and the Omron R7 (HEM 637-IT). Devices and Technology. *Blood Pressure Monitoring*, *11*, 165-171.

Trimmel, M., Meixner-Pendleton, M., & Haring, S. (2003). Stress Response Cause by System Response Time when Searching for Information on the Internet. *Human Factors, 45*, 615-621.

Truong, K.P, Van Leeuwen, D.A., and Neerincx, M.A. (2007). Unobtrusive Multimodal Emotion Detection in Adaptive Interfaces: Speech and Facial Expressions. In Schmorrow & Reeves (Eds.) *Foundations of Augmented Cognition*, 3rd Edition, LNAI 4565 proceedings (pp. 354-363).

Ursin, H., Baade, E., & Levine, S. (Eds.) (1978). *Psychobiology of Stress: A Study of Coping Men.* New York: Academic Press.

Ursin, H., Eriksen, H.R. (2004). The cognitive activation theory of stress. *Psychoneuroendocrinology*, *29*, 567-592.

Van Cauter, E., Leproult, R., & Kupfer, D.J. (1996). Effects of gender and age on the levels and circadian rhythmicity of plasma cortisol. *Journal of Clinical Endocrinology & Metabolism, 81*, 2468-2473.

Van Dale. (2003). *Van Dale Groot woordenboek der Nederlandse taal* (versie 1.3 Plus op basis van de dertiende uitgave). Utrecht/Antwerpen: Van Dale Lexicografie BV.

Van Kuijk, J. (2005). *Near Field Communication: very touching*. Retrieved October 23, 2006, from http://uselog.blogspot.com/2005_02_01_uselog_archive.html

Virzi, R.A., Sokolov, J.L., & Karis, D. (1996). Usability Problem Identification Using Both Low- and High-Fidelity Prototypes. *CHI 96, April 13-18.*

VIVA. *Viva Forum*. Retrieved January, 2003, from http://forum.viva.nl/

Welford, A.T. (1985). Changes of performance with age. In N. Charness (Ed.) *Aging and Human Performance.* New York: Wiley.

Wickens, C.D., Lee, J.D., Liu, Y., & Gordon-Becker, S. (2004). *Introduction to Human Factors Engineering* (2^{nd} ed.). Upper Saddle River, NJ: Pearson Education.

Wiedenfeld, S.A., O'Leary, A., Bandura, A., Brown, S., Levine, S., & Raska, K. (1990). Impact of Perceived Self-Efficacy in Coping With Stressors on Components of the Immune Sytem. *Journal of Personality and Social Psychology*, *59*, 1082-1094.

Xie, B. (2002). Older adults, computers and the Internet: Future directions. *Gerontechnology, 2,* 289-305.

Yerkes, R.M., & Dodson, J.D. (1908). The Relation of Strength of Stimulus to Rapidity of Habit-Formation. *Journal of Comparative Neurology and Psychology, 18*, 459-482.

Appendix

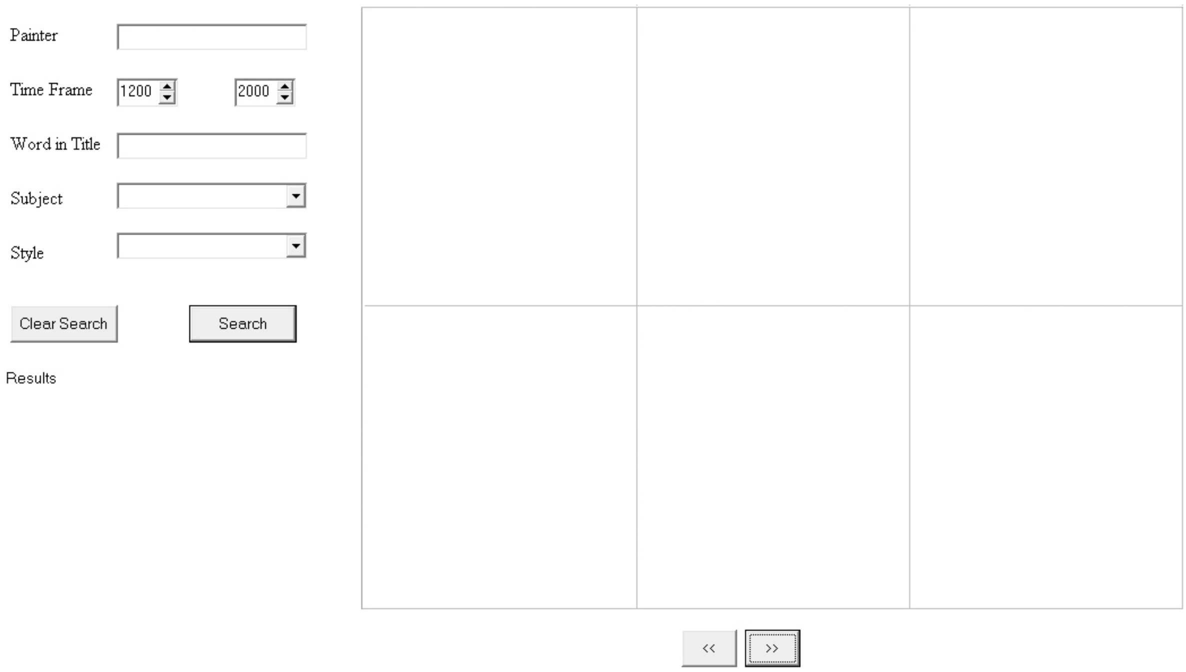

Figure A.1: Screen capture of the first art image retrieval prototype after a searchword is filled in, or a selection is made at the time frame or drop-down menu's and 'Search' is pressed. The art images appear in squares on the right of the search interface. No art images are shown due to possible copyright issues. The two buttons with the arrows appear under the images. With these buttons a user can browse forward or backward through the art images. The number of results is shown under 'Clear Search'.

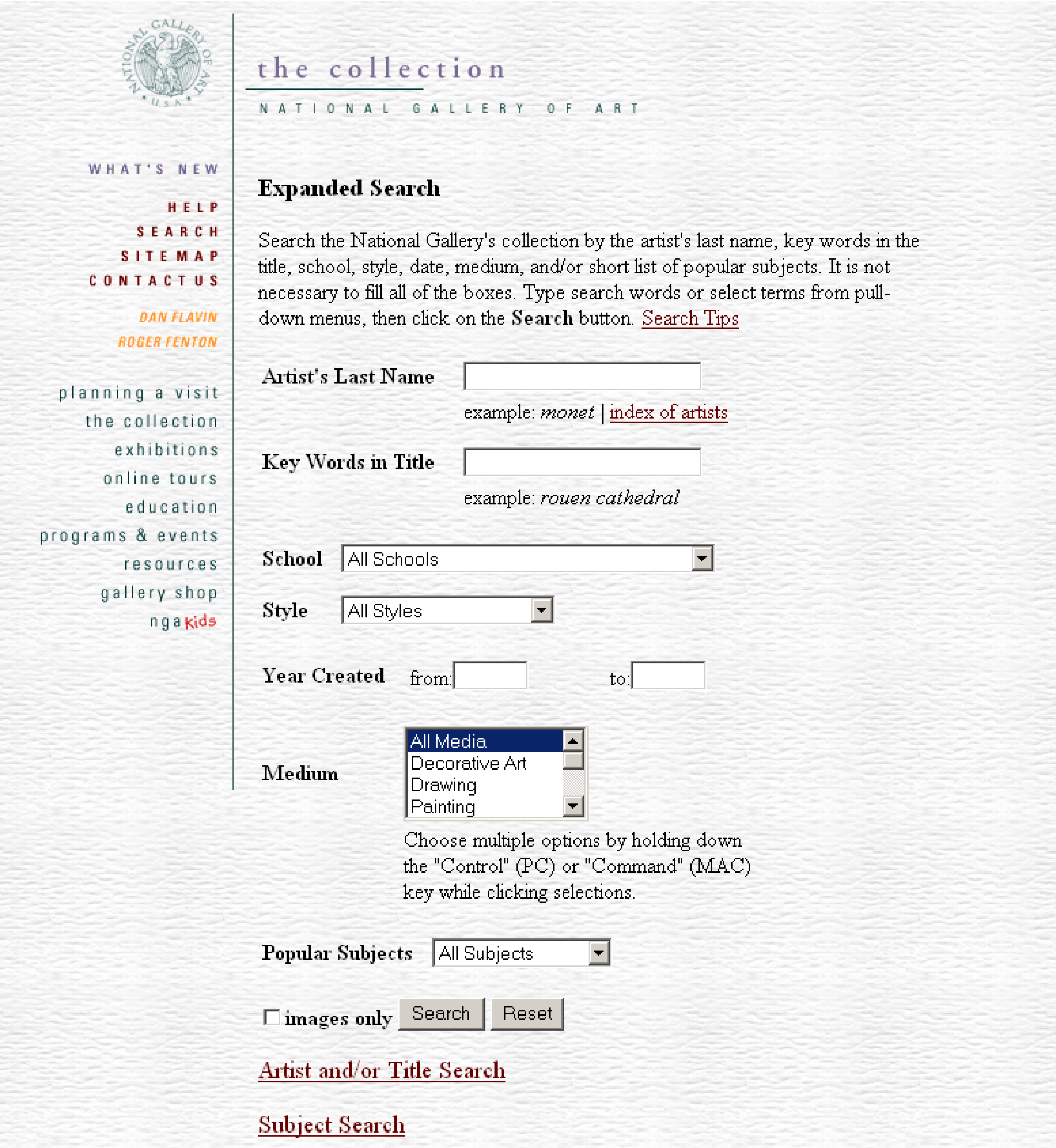

Figure A.2: Screen capture of the National Gallery of Art search interface (NGA, 2004)

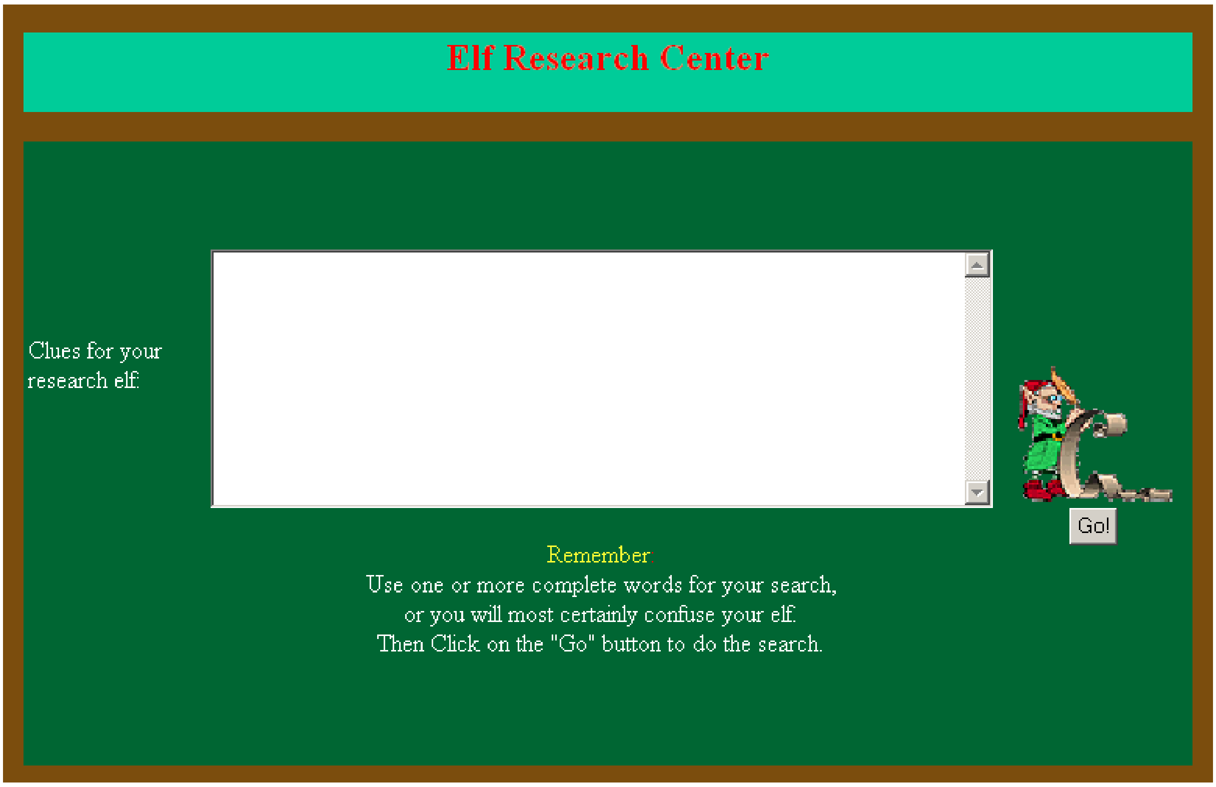

Figure A.3: Screen capture of the Tigertail Virtual Museum search interface (TVM, 2004)

Figure A.4: Screen capture of the Help system after a searchword is filled in, or a selection is made at the drop-down menu's and 'Zoek' (Search) is pressed. Art images appear in squares on the right of the search interface. No art images are shown due to possible copyright issues. The number of results is shown in blue, under 'Wissen' (Clear) at 'Resultaten' (Results). Under the art images, a 'Volgende' (Next) button and 'Vorige' (Previous) button appears. With these two buttons a user can browse through the results.

Summary

Stress in Senior Computer Interaction

Seniors have been found to face difficulties learning and using new media. To overcome problems with accessibility, digital information systems should be engineered with reference to the knowledge and capabilities of seniors, but there is more to satisfying interaction; low stress. Little is known about the mechanisms behind stress in senior computer interaction. In this thesis, the complexity of a system and the expectancies seniors have about their performance were considered as factors that could influence acute stress. Six experiments focused on improving the accessibility of digital information systems and the experiments contributed to understanding stress in senior computer interaction. The first five experiments were conducted within the art image retrieval domain and the last experiment within the Alzheimer text retrieval domain.

The first three experimental evaluations reported in Chapter 2 showed how to design an accessible retrieval system for seniors. It was demonstrated that young experienced users benefited from a cognitively engineered image retrieval system, while seniors with less experience not. The seniors showed high levels of stress. Cognitive engineering is an iterative process, and for the senior computer users extra steps were necessary to provide high accessibility. An error analysis shed light on the problems the seniors faced during interaction. Three different redesigns were constructed to overcome the problems they faced, and in an experiment it was shown that all three redesigns provided seniors with high accessibility and low stress. Although this is a valuable improvement, it was unclear why the seniors were stressed during the first experimental evaluation.

In Chapter 3, an experiment is reported that aimed to examine the effect of system complexity on stress in senior computer interaction. It was expected that when seniors interact with a high complexity system that their acute subjective- and physiological stress is higher than when they interact with a low complexity system. The results of the experiment showed no significant effect of system complexity on acute physiological- and subjective stress.

In Chapter 4, the effect of complexity and expected task performance on stress was studied. Expected task performance was manipulated by performance feedback information, and it was expected that acute subjective- and physiological stress is higher when seniors receive negative performance feedback information than when they receive positive feedback information. Again, system complexity had no effect on acute stress. It was found that when seniors receive negative performance feedback, expected task performance decreased and subjective stress increased. When seniors interact with a high complexity system, their expected task performance declines while when they interact with a low complexity system

expected task performance increases. Seniors were significantly more efficient and effective in performing search tasks on the low complexity system than on the high complexity system.

Similar results were found in an experiment reported in Chapter 5, which was conducted within the Alzheimer domain. In addition, it was shown that inexperienced seniors benefit from a touchscreen and ABCDE keyboard as input devices to retrieve digital information. An interesting finding of the experiment was that many seniors did not trust the Alzheimer information that was presented by the Alzheimer text retrieval system, which could have been caused by the minimal presentation of the information or an exposure of mental discomfort and denial.

Stress plays a role in senior computer interaction. Acute subjective stress increases when seniors receive negative performance feedback information. The complexity of a system has no effect on subjective stress, but it does have an effect on the expectancies seniors have about their future task performance and the efficiency and effectiveness tasks are performed. The knowledge that was gathered from the experiments can be used as requirements for the design of accessible systems for seniors, e.g., it is important to strive for large fonts and buttons, native language and positive feedback information in the design. By iterative design and by taking the experience and abilities of seniors into account, most seniors, even seniors without any computer experience, are able to successfully interact with a digital information system.

Samenvatting

Stress in Senior Computer Interactie

Senioren kunnen problemen ondervinden in het leren en gebruiken van nieuwe media. Om deze problemen te kunnen ondervangen moet er in het ontwerp van digitale informatiesystemen rekening gehouden worden met de kennis en mogelijkheden van senioren. Maar er is meer om succesvol te kunnen iterageren; weinig stress. Er is weinig bekend over de stress die ouderen ervaren tijdens computergebruik. In dit proefschrift is onderzocht of de complexiteit van een systeem en de prestatieverwachting van senioren van invloed zouden kunnen zijn op acute stress. Zes experimenten richtten zich op het verbeteren van de toegankelijkheid van digitale informatie systemen en de experimenten hebben een bijdrage geleverd aan kennis over stress in senior computer interactie. In de eerste vijf experimenten werd een kunstplaatjeszoeksysteeem gebruikt als onderzoeksplatform en in het laatste experiment werd een zoeksysteem gebruikt waarmee naar tekstuele informatie over de ziekte van Alzheimer gezocht kon worden.

De eerste drie experimentele evaluaties in hoofdstuk 2 toonden aan hoe een toegankelijk zoeksysteem ontworpen kan worden voor senioren. Jongeren profiteerden van een zoeksysteem dat ontworpen was volgens 'cognitive engineering' methode, terwijl senioren met minder computerervaring er niet van profiteerden. Uit observaties bleek dat de senioren gestrest waren tijdens het experiment. Cognitive engineering is een iteratief proces en voor de senioren waren er extra ontwerpstappen nodig om de toegankelijkheid van het systeem te vergroten. Door het uitvoeren van een foutenanalyse werd duidelijk welke problemen de senioren ondervonden. Drie verschillende herontwerpen werden geconstrueerd om de problemen die de senioren ondervonden te overkomen. In een experiment bleek dat deze herontwerpen de toegankelijkheid voor senioren vergroten. Ondanks dat dit een waardevolle verbetering is, was het onduidelijk waarom de senioren gestrest waren in het eerste experiment.

In Hoofdstuk 3 wordt een experiment beschreven waarin het effect van systeem complexiteit op acute stress in senioren is onderzocht. De verwachting was dat senioren subjectief- en fysiologisch meer gestrest zijn als ze interageren met een systeem van hoge complexiteit dan met een systeem van lage complexiteit. De resultaten van het experiment lieten zien dat er geen effect is van de complexiteit van een systeem op acute subjectieve- en fysiologische stress.

In Hoofdstuk 4 werd het effect van complexiteit en verwachte taakprestatie op stress onderzocht. Verwachte taakprestatie werd gemanipuleerd door prestatiefeedback en de verwachting was dat acute subjectieve- en fysiologische stress hoger is als senioren negatieve feedback krijgen dan als ze positieve feedback krijgen. De complexiteit van een systeem had geen effect op acute stress. Als senioren negatieve feedback kregen, ging hun verwachte taakprestatie omlaag en hun subjectieve stress ervaring omhoog.

Als senioren met een systeem van hoge complexiteit interageerden ging verwachte taakprestatie omlaag, terwijl als ze met een systeem van lage complexiteit interageerden ging verwachte taakprestatie omhoog. Senioren waren significant efficiënter en effectiever in het uitvoeren van zoektaken op het systeem van lage complexiteit in vergelijking tot het systeem van hoge complexiteit.

In het experiment dat in hoofdstuk 5 is beschreven, dat zich binnen het Alzheimerdomein afspeelde, zijn soortgelijke resultaten gevonden. Bovendien werd er gevonden dat onervaren senioren profiteerden van een touchscreen en ABCDE keyboard als inputapparaat om digitale informatie te zoeken. Een interessante vinding was dat senioren de Alzheimerinformatie, dat gepresenteerd werd door het Alzheimerzoeksysteem, niet vertrouwden. Dit werd mogelijk veroorzaakt door de minimale presentatie van de informatie of het was een teken van mentale dissonantie en ontkenning.

Stress speelt een rol in senior computer interactie. Acute subjectieve stress gaat omhoog als senioren negatieve prestatiefeedback krijgen. De complexiteit van een systeem heeft geen effect op subjectieve stress, maar het heeft wel een effect op de verwachtingen die senioren hebben over hun toekomstige taakprestatie en de efficiëntie en effectiviteit waarmee de taken worden uitgevoerd. De kennis die verzameld is uit de experimenten kan worden gebruikt als systeem- en designeisen voor het ontwerp van toegankelijke systemen voor senioren. Het is onder andere belangrijk om in het ontwerp te streven naar grote fonts en knoppen, moedertaal en positieve feedback informatie. Door iteratief ontwerp en door rekening te houden met de ervaring en capaciteiten van senioren in het ontwerp, kunnen de meeste senioren zonder enige computerervaring succesvol interageren met een digitaal informatiesysteem.

Curriculum Vitae

1978	Born in Hardenberg, The Netherlands
1990-1995	HAVO, RSG, Coevorden
1995-1996	High School, Barr-Reeve Montgomery, Indiana (US)
1996-1997	VWO, Luzac College, Zwolle
1997-2002	Psychology - Cognitive Ergonomics, Utrecht University
2002-2008	Ph.D. student at the J.F. Schouten School for User-System Interaction, Eindhoven University of Technology