File ID uvapub:81827
Filename Thesis
Version unknown

SOURCE (OR PART OF THE FOLLOWING SOURCE):
Type PhD thesis
Title Assessing acceptance of assistive social robots by aging adults
Author(s) M. Heerink
Faculty FNWI: Informatics Institute (II)
Year 2010

FULL BIBLIOGRAPHIC DETAILS:
http://hdl.handle.net/11245/1.327918

Assessing acceptance of assistive social robots by aging adults

Marcel Heerink

Assessing acceptance of assistive social robots by aging adults Marcel Heerink

Assessing acceptance of assistive social robots by aging adults

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula der Universiteit
op woensdag 3 november 2010, te 10.00 uur

door Marcel Heerink

geboren te Rheden

Promotor: prof. dr. B.J. Wielinga

Co-promotores: dr. ir. B.J.A. Kröse
dr. V. Evers

Overige leden: prof. dr. K. Dautenhahn
prof. dr. ir. F.C.A. Groen
prof. dr. L. Hardman
prof. dr. M. A. Neerincx
prof. dr. L. P. de Witte

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

This research was supported in part by the Hogeschool van Amsterdam, University of Applied Sciences (HvA) and in part by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020 (Cogniron).

Acknowledgements

This thesis and the underlying research would not have been possible without the support and friendship of many people and institutions. First of all, I worked with three fascinating and inspiring people: Bob Wielinga, Ben Kröse and Vanessa Evers. They guided me in many ways, the past few years.

Furthermore, I would like to express my gratitude to the staff and caretakers as well as the test participants of De Kiekendief, De Overloop and De Archipel in Almere, De Uiterton in Lelystad and De Emtinckhof in Loosdrecht for their trust and cooperation. Also, I am extremely grateful for those participants who were willing to receive us at their homes and let us mess with their computers.

Also I would like to thank Wouter van Gils, Willem Notten, Bas Terwijn and Rogier Voors for their work on programming the iCat system, Nicolai Sequeira Geraldes and Tormo Antonides for their work on Annie and Susanne Roelofs, Hatice Çal, Rick van Midde, Albert van Breemen and Martin Saerbeck for their support. In addition, I like to thank my colleagues at the Information Engineering Institute and the people at HCS for their encouragement. And I thank Tonnie Triezenberg for making the whole project possible in the first place.

Moreover I wish to thank Ruud Smeulders, Rabobank Foundation, Netrex and Fons Kuijk of CWI for enabling us to work with Steffie and develop Annie, and Frederico Pecora for the RoboCare materials.

Furthermore, I am eternally gratefull to Irma van der Puijl for helping me make this a readable text and to Heidi, Lennart, Wietse and Eline for their support and patience.

Contents

Publications

Parts of this manuscript have been published in the following refereed journal papers:

- Heerink, M., B.J.A. Kröse, B.J. Wielinga, and V. Evers, Measuring acceptance of assistive social agent technology by older adults: the Almere model. *International Journal of Social Robotics*, 2010. 2(3), (Accepted)
- Heerink, M., B.J.A. Kröse, V. Evers, and B.J. Wielinga, Relating conversational expressiveness to social presence and acceptance of an assistive social robot. *Virtual Reality*, 2010 14(1): p. 77-84
- Broekens, J., M. Heerink, and H. Rosendal, *The effectiveness of assistive social robots in elderly care: a review*. Gerontechnology journal, 2009. 8(2): p. 94-103.
- Heerink, M., B. Kröse, B. Wielinga, and V. Evers, *Influence of Social Presence on Acceptance of an Assistive Social Robot and Screen Agent by Elderly Users*. Advanced Robotics, 2009. 23(14): p. 1909-1923.
- Heerink, M., B.J.A. Kröse, B.J. Wielinga, and V. Evers, The Influence of Social Presence on Acceptance of a Companion Robot by Older People. *Journal of Physical Agents – Special Issue on Human interaction with domestic robots*, 2008. 2(2): p. 33-40.
- Heerink, M., B. Kröse, V. Evers, and B. Wielinga, Studying the acceptance of a robotic agent by elderly users. *International Journal of Assistive Robotics and Mechatronics*, 2006. 7(3): p. 33-43.

Refereed conference papers:

- Heerink, M., B. Kröse, B. Wielinga, and V. Evers. Measuring acceptance of an assistive social robot: a suggested toolkit. in Ro-man 2009 – In *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication 2009*. Toyama, Japan, p. 528-533
- Heerink, M., B. Kröse, B. Wielinga, and V. Evers. Measuring the influence of social abilities on acceptance of an interface robot and a screen agent by elderly users. In *Proceedings HCI 2009*. Cambridge, p. 430-439
- Heerink, M., B.J.A. Kröse, B.J. Wielinga, and V. Evers. The influence of perceived adaptiveness of a social agent on acceptance by elderly users. In ISG 2008 - *The 6th International Conference of the International Society for Gerontechnology*. 2008. Pisa, p. 57-61
- Heerink, M., B.J.A. Kröse, B.J. Wielinga, and V. Evers. Measuring Perceived Adaptiveness in a robotic eldercare companion. In *HRI 08 - Workshop Robotic Helpers*. 2008, p. 19-21
- Heerink, M., B.J.A. Kröse, B.J. Wielinga, and V. Evers, Enjoyment, Intention to Use and Actual Use of a Conversational Robot by Elderly People, In

Proceedings of the third ACM/IEEE - International Conference on Human-Robot Interaction (HRI). 2008, Amsterdam, p. 113-120.
- Heerink, M., B. J. A. Kröse, B. J. Wielinga and V. Evers. The Influence of Social Presence on Enjoyment and Intention to Use of a Robot and Screen Agent by Elderly Users. *Proceedings RO-MAN 2008*, München, p. 695-700
- Heerink, M., B.J.A. Kröse, B.J. Wielinga, and V. Evers. *Observing conversational expressiveness of elderly users interacting with a robot and screen agent. In Proceedings International Conference on Rehabilitation Robotics (ICORR).* 2007. Noordwijk, p.751-756
- Heerink, M., B.J.A. Kröse, B.J. Wielinga, and V. Evers. The Influence of a Robot's Social Abilities on Acceptance by Elderly Users. In *Proceedings RO-MAN. 2006*. Hertfordshire, 521-526
- Heerink, M., B. Kröse, B. Wielinga, and V. Evers. Human-Robot User Studies in Eldercare: Lessons Learned. In *Smart Homes And Beyond: ICOST2006: 4th International Conference on Smart Homes and Health Telematics*. 2006, p. 31-38.

Refereed video publications based on this research:

- Heerink, M., B.J.A. Kröse, B.J. Wielinga, and V. Evers. iCat in Eldercare, In C. Bartneck & T. Kanda (Eds.) HRI Caught on Film. In *Proceedings of the 2nd ACM/IEEE International Conference on Human-Robot Interaction*, Washington DC, march 2007.
- Heerink, M., B.J.A. Kröse, B.J. Wielinga, and V. Evers. Video: Responses to a social robot by elderly users. In IEEE/RSJ 2008 *International Conference on Intelligent Robots and Systems.* 2008. Nice.

1. Introduction

He that will not apply new remedies must expect new evils, for time is the greatest innovator.

Francis Bacon

1.1 Assessing acceptance…

The research described in this thesis is about the acceptance of social robots which can help older adults socially, mentally and physically. More precisely, its goal is to find a way to measure how willing these older adults are to make these systems a part of their daily lives. There is a methodology that is often used to measure the willingness to use a certain technology - it is called Technology Acceptance Modeling. It has not been developed to be used for this specific technology and user group, but we will establish how we can adapt it to be usable. This means that at the end of this thesis, we present a technology acceptance methodology that can be applied to social robots, used to assist older adults.

First, in this introductory chapter we will describe the motivation of this study by discussing the implications of a current demographic shift, the role robots can play in the near future and the relevance of studying technology acceptance in this context. Next, we will set the focus by defining our key concepts and specifying our goal and research questions. Subsequently, we will describe the approach of our research and overview the setup of this thesis.

1.2 Motivation: a demographic shift

As the baby boom generation is aging, the number of elderly citizens is growing considerably and expected to grow even more over the coming decades – not only in the Netherlands (see Figure 1.1) but throughout the entire industrialized world (Soede et al. 2004; Verzijden and Fransen 2004). In 2006 it was estimated that at that moment 15% of the total population of industrialized countries was over 60 years old and it was projected to be around 30% in the year 2040 (United Nations 2007).

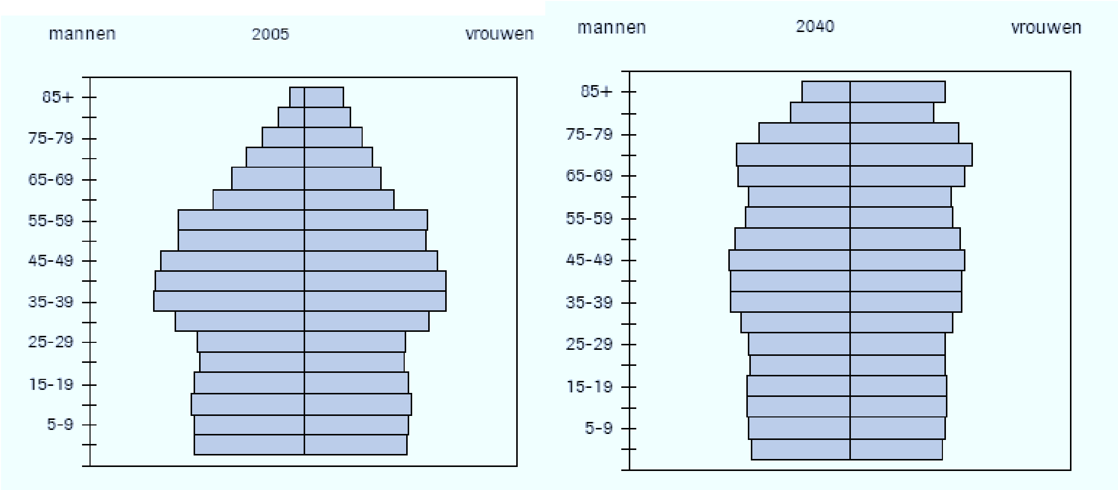


Figure 1.1. Demographic developments in the Netherlands (source: Verzijden & Fransen, 2004)

This development is even more alarming if we realize that alongside this growth in the elderly population, in the same industrialized countries we face short- and long-term labor shortages, especially in the healthcare sector. This is already showing in the first decades of the 21st century and it will aggravate dramatically in the twenties and thirties. Studies that describe these shortages (Strunk et al. 2001; Barea et al. 2004) also point out a related development, concerning the *costs* in the health-care sector for elderly in the industrialized countries. Since the mid 1990's, nursing home costs have more than doubled, which of course poses pressure on the staffing budget.

Clearly the unprecedented increase of the elderly population along with increased labor shortage and the explosion of costs pose extreme challenges to society. Efforts to meet these challenges include projects that explore the applicability of technological advances like intelligent systems that enable elderly people to live independently (Sixsmith 2002; Stanford 2002; Consolvo et al. 2004; Cesta and Pecora 2005). This research is partly motivated by the notion of staff relief: in intelligent environments, elderly citizens would need less human assistance or need human assistance at a much later stage in their aging process. However, also many research projects are simply motivated by the desire to make our old age more comfortable (Bahadori et al. 2003; Miller et al. 2004; Pollack 2005).

In general, we see that research concerning the use of intelligent technology for older adults addresses their needs in three fields: physical, cognitive and socially. Some examples:

- the physical needs are met with solutions like intelligent wheelchairs, walking aids, exoskeletons and robotic butlers (Graf 2001; Yanco 2001; Guizzo and Goldstein 2005);
- the cognitive needs are addressed with monitoring systems and adaptive reminder devices (Russell et al. 2006; Scanaill et al. 2006);
- the social needs are met with communication devices and robotic pets

(Beck 2003; DiSalvo et al. 2003).

Apart from projects that address one of these fields, there are projects that address all three categories simultaneously. They generally concern different types of automated homes that provide a caring environment, adapting to the changing needs of aging adults (Giuliani et al. 2005; Jung et al. 2005; Hurst et al. 2006; Cesta et al. 2007). This typically includes a monitoring system, central control of household devices, medication assistance, robotic devices to assist when bathing or toileting, and household maintenance. In such homes, professional human assistance is only needed in times of emergencies.

These numerous projects mark the dawn of assistive technology that will become part of the lives of aging citizens. Part of this development is a specific technology that is, due to its easiness of use and multiple possibilities, very suitable to be applied to this specific context: assistive social robots. In this thesis we focus on this technology.

1.3 Assistive companions

Projects concerning the development of assistive robots for older adults address one or more of the three above mentioned types of needs, either as independent devices or as a part of intelligent environments. As independent devices, they address physical and cognitive needs by supporting mobility, basic activities like eating, bathing, toileting, getting dressed, providing household maintenance, monitoring and maintaining safety (Mynatt et al. 2000; Bickmore 2004; Taggart et al. 2005; Decker 2008). As part of an intelligent environment, they usually are envisioned to function as interfaces, facilitating communication between the environment and its user (Jung et al. 2005; Cesta et al. 2007).

Focusing on social needs, some studies demonstrate how robots can have the ability to provide (pet like) companionship (Friedman et al. 2003; Wada et al. 2003). This demonstrates how robots can anticipate the need for a social entity by the user to build an emotional relationship with. However, this possibility to build an emotional relationship, combined with the easiness of use of an interface that is controlled by social interaction, not only responds to social needs. It also increases acceptability, as several studies show (Bickmore 2004; Bickmore et al. 2005; Wu and Miller 2005).

This social engagement is not something exclusively for robots. In general, findings from ‘affective computing’ show that for many user groups applications are easier accepted and enjoyed more if emotional features are added to the interaction with computers. It has been established that both the recognition of emotions and the appearance of emotional features in interaction with computer applications turn out to facilitate and enrich communication as perceived by the users (Picard 1997; Picard and Daily 2005). If a tutorial agent for example shows regrets if a mistake is made and joy if a task is performed correctly, this makes the application more acceptable (Elliott et al. 1997); if a system shows empathy in its feedback, it makes the users appreciate the system more and they work with it longer and more intensively (Conati et al. 2005).

The zoomorphic or anthropomorphic embodiment of robots that communicate socially can strengthen this effect of social interaction. It makes them more believable as social actors and often gives them more possibilities to communicate socially, like using facial expressions or gestures. This implies that robots can be very acceptable and enjoyable as social actors and facilitate interaction with advanced technology (for extensive arguments see Wilkes et al. (1997), Breazeal (2003) and Nijholt (2003)). This is a great advantage which makes it possible for robots to be assistive devices that are easy to communicate with, besides their abilities to meet the different needs of older adults.

1.4 The issue of acceptance

We can thus conclude that for assistive technology in general, robots have a high potential to play a role in eldercare in the (near) future, not only by realizing staff relief but also by improving the quality of eldercare, providing services that are beyond human staff capabilities (Baltus et al. 2000; Barea et al. 2004; Forlizzi et al. 2004). Still, to be of use at all, older adults have to be willing to actually use this type of technology: robots have to be accepted. Several research projects concerning assistive technology show that, although a large category of elders may be open to assistive technologies, technology acceptance remains a delicate matter. There is a range of systems and applications developed to fit the demands of elderly that are still not actually being used, often because of factors like stigmatization, (non-)adaptiveness of the device, or social influences ('peer pressure'). Consequently, there is a need for sophisticated strategies to develop technology that will actually be adopted (Forlizzi et al. 2004; Bickmore et al. 2005; Kidd et al. 2006; Cesta et al. 2007).

Several aspects of acceptance have been studied, but in general, studies on robot acceptance, especially those on social intelligence in human-agent interaction concerning elderly people, are based on either theoretical considerations or qualitative findings from a small set of users (Forlizzi et al. 2004; Graf et al. 2004; Kaplan 2004; Libin and Cohen-Mansfield 2004; Bickmore 2005; Taggart et al. 2005). Some studies on acceptance of technologies that are somewhat similar (assistive technology, socially communicative systems) apply technology acceptance modeling (TAM). This is a quantitative approach to identify influences on the intention of a population to make use of a specific technology and on the actual usage (Van der Heijden 2004; Wilson and Lankton 2004; Wu et al. 2005; Chesney 2006). An acceptance model can predict how well a system will be accepted by a certain user group and explain differences between individuals or sub groups. In addition, it can be used to research differences between different system types or conditions.

However, applying TAM to evaluate acceptance of assistive technology or entertainment systems is still quite different from using it to evaluate acceptance of systems that can be perceived as a social entity, such as a robot or screen agent. Moreover, acceptance modeling has not yet been applied specifically to elderly users. There is also no specific model for either robots or elderly users: influences that are known to be of importance when it comes to acceptance of a social entity have never been included by any technology acceptance model and neither have influences that are known to affect elderly users.

This might be due to the focus on development (mainly on design and functionality) when evaluating these systems. If we look at publications concerning HRI surveys and HRI benchmarks (e.g. Yanco and Drury 2004; Kahn et al. 2006; Feil-Seifer et al. 2007), we see that only a survey by Goodrich and Schultz (2007) leaves some room for focus on user acceptance, but identifying or measuring influences is not discussed.

Therefore, we can state that it would be a relatively novel approach to use technology acceptance methodology specifically to address the acceptance of assistive social robots by elderly users. In the next two sections we will specify our goal, problem statement, research questions and approach, followed by an overview of the content of this thesis.

1.5 Focus of this study

The aim of this study is to develop a model that can be used to explain and predict influences on acceptance of robots by elderly users. In order to develop this instrument, we want to identify these influences by developing a model and methodology that can be used to predict and explain acceptance of an assistive social robot by elderly users.

We focus on 'assistive social robots': a type of robot that communicates socially and assists older adults. In the next chapter we will specify the applicable systems with a survey of assistive social robots used in eldercare, but to establish our focus, we need to further specify the term *assistive social robot*.

First, we define an *assistive* robot as *one that gives aid or support to a human user*. This means it is designed to give information or to perform activities with which it helps someone. This does not necessarily have to do with physical or mental disabilities: a robot which gives directions can yet be called assistive.

Secondly, to be more specific on the term *social* in the context of assistive robots, we can make use of several overlapping definitions of the concepts of social abilities, sociability and social intelligence (Kihlstrom and Cantor 2000). Since the systems described by Fong et al. (Fong et al. 2003) and Breazeal (Breazeal 2003) are actually robots that are potentially assistive, we will adopt their understanding of socially interactive robots: *those that people apply a social model to in order to interact with and understand*. This means people use their understanding of social interaction both to interpret anything the robot communicates and to communicate with it. Interaction is carried out by providing a 'natural' interface by employing human-like social cues and communication modalities. This means that some kind of social (anthropomorphic or zoomorphic) behavior is at least a part of the interaction, but interaction does not necessarily have to be limited to this (as is the case with socially assistive robots).

Concerning the sequence of the words *assistive* and *social*, we found that our understanding of the researched category essentially differs from the category of that *socially assistive* robots as defined by Feil-Seifer et al. (Feil-Seifer and Mataric 2005; Feil-Seifer et al. 2007). The latter states that *socially assistive* robots are an intersection of assistive robots (AR) and socially interactive robots (SIR). For Feil-Seifer et al. *socially assistive* means that assistance is provided through social interaction (like pet-like companionship). This specification rather limits the scope, for many social robots are physically assistive while their social interaction is not meant to be of social assistance (see for example Graf, Hans, & Schraft, 2004 and Mukai et al. 2008). For that reason we will not use the term *socially assistive* robots, but *assistive social* robots: they are not necessarily socially assistive, but generally assistive and socially interactive.

Thus, we come to the following definition: *assistive social robots are socially interactive robots that are in some way assistive to their user group*. Within the present study this user group consists of older (65+) adults. In chapter 2 we will further specify the systems that belong to this category, give examples of robots that are either social or assistive and survey publications on assistive social robots that have been developed for or applied to older adults.

Within this study, a *robot* can be any system to which the general understanding of a robot applies, as defined in the Encyclopaedia Britannica: 'any automatically operated machine that replaces human effort, though it may not resemble human beings in appearance or perform functions in a humanlike manner' (Encyclopaedia Britannica Online 2010). In several sections we will, however, also include findings concerning screen personalities (also called screen agents or relational agents), because research comparing robots and agents generally shows that people respond to them in a similar way (Shinozawa et al. 2003; Bartneck et al. 2004; Shinozawa et al. 2005).

Furthermore, we need to operationally define the term *acceptance*, for there are multiple interpretations (Dillon 2001; Tscheligi and Bernhaupt 2004; Wu et al. 2005). We use the term according to its meaning within the field of technology acceptance modeling (Lee et al. 2003): *acceptance means that a system or application is actually used after an initial exploration* (in section 2.4 we will introduce and specify the term *use*) and a higher degree of acceptance means there is a more frequent use of the specific technology.

To arrive at our aim to explain and predict influences on acceptance of robots by elderly users, we first want to know whether there is a model that we can use, perhaps with just minor modifications. If this is not the case, we want to know what influences should be incorporated that are not already represented in any model. And of course we want to know whether the model is appropriate if we incorporate these influences.

Our main research question is the following:

> *How can we explain and predict the influences on acceptance of assistive social robots by elderly users?*

To answer this question we need to answer the following sub questions:

1. To what extent is the most prevailing technology acceptance model able to explain and predict acceptance of assistive social robots by elderly users?
2. What evidence can be found concerning alternative influences on acceptance of assistive social robots by elderly users?
3. If the most prevailing model is not able to adequately explain and predict acceptance of assistive social robots by elderly users, can we set up a new model by incorporating new influences and prove it to have a better explanatory power than that model?

Answering the first and third question means that a model should meet the following criteria:

1. It should have the ability to explain acceptance under a wide variety of experimental conditions.
2. It should show robustness during quantitative analysis.
3. It should aim to identify the main influences on acceptance of assistive social robots by elderly users.

The broad explanatory power of various models will be analyzed, using quantitative instruments and standard statistical procedures (these will be discussed in section 2.4.4).

1.6 Approach and outline

To answer the first question, we will explore technology acceptance models. In chapter 2 we will describe this field. We will give an overview of research done with several models and arrive at the UTAUT model, which claims to incorporate all relevant influences on technology acceptance in general. In this chapter we will also present an inventory of assistive social robots that are either developed for or applied to elderly users. We will summarize and analyze the user studies that are published regarding these robots, and establish their relation to our research.

In chapter 3 we will describe our experiments, using two different systems (a robot and a screen agent) to test a preliminary model based on UTAUT. We will present our findings and evaluate the applicability of this model in this particular context, thus answering the first question.

As the conclusion will be that the model needs further development, in chapter 4 we will identify influences that have not been part of the UTAUT based model and should be represented according to findings in related research. These will concern both influences that have been part of related acceptance models and influences that are new to acceptance methodology.

To prove this newly developed model appropriately, we will describe our validation experiments on different types of assistive social agents in chapter 5, 6 and 7. In chapters 5 and 6 these experiments are focused on establishing the validity of the constructs that have not yet been part of any related acceptance model and in chapter 7 the focus will be on establishing the validity of the model (in two experiments that include a usage period).

In chapter 8 we will use the collected data to refine our model and methodology. We will both establish confirmation of the model as it has been developed so far and use explorative statistical techniques to improve its predictive strength.

Chapter 9 will show a summary of our findings. We evaluate them in relation to the research questions and establish the main contributions of this research. In the final section of this chapter we will conclude this thesis by a discussion which includes looking forward to future research.

2. Robots and acceptance by elderly users

Anything that is in the world when you're born is normal and ordinary and is just a natural part of the way the world works. Anything that's invented between when you're fifteen and thirty-five is new and exciting and revolutionary and you can probably get a career in it. Anything invented after you're thirty-five is against the natural order of things.

Douglas Adams

Parts of this chapter have been published earlier in (Broekens et al. 2009; Heerink et al. 2010b)

2.1 Introduction

In this chapter we provide an overview of research and theories related to acceptance of assistive social robots, which is our focus as we have set out in Chapter 1. After a categorization of robots which establishes the position of assistive social robots in robot technology, we will list the robots that have been developed for or have been applied to elderly users. We will establish whether they have been subject to user studies and we will briefly discuss the results of these studies and how they relate to our research. Furthermore we explore the field of technology acceptance, looking at its history and recent developments, and its application to robots. At the end of this chapter we summarize our findings and point out how the presented research is determining our project.

2.2 Social and assistive robots

In the previous chapter we discussed our notion of the concepts 'social' and 'assistive' within the context of robots used by older adults. We will now deepen our understanding of these concepts and show examples of developments of both robot types. Subsequently, we will provide a categorization which clarifies the position of assistive social robots within the field of human-robot interaction.

In section 1.5 we defined our understanding of social robots by referring to Fong et al. (Fong et al. 2003): *social robots are those that people apply a social model to in order to interact with and understand*. To evoke this application of a social model, a robot has to have certain characteristics – it has to look or behave in a way that suggests that applying a social model does make sense. In general, this is the case if a robot exhibits one or more of the characteristics shown in the following seven point list:

1. Express and/or perceive emotions (emotions can be expressed by facial expressions, by sounds or by speech - they can be perceived for example by analysis of facial expressions or sounds).
2. Communicate with high level dialogue (e.g. the abilities to ask questions and to use dialogue to solve problems mutually).
3. Learn/recognize models of other agents.
4. Establish/maintain social relationships.
5. Use natural cues (e.g. gaze, gestures).
6. Exhibit distinctive personality and character.
7. May learn/develop social competencies.

This list contains *characteristics* that can be made operational by implementing social *abilities* in robots. These abilities can be scalable and can make a robot more or less sociable. For example, the characteristic of expressing emotions can be implemented by enabling a robot to use facial expressions. We can assign a higher degree of expressiveness to a robot that has either more possibilities for facial expressions or is able to use them more sophistically than another one with fewer possibilities or less sophistication in applying them. Both robots would then be characterized as social robots with the characteristic of being expressive, although there are obvious differences concerning their social abilities. In chapter three we will work out a set of actual social abilities in order to implement these in two different robotic systems.

Examples of robots that are designed upon these principles and that represent milestones in the field of human-robot interaction are Kismet and museum robot Sage (see Figure 2.1). Kismet has been developed at Massachusetts Institute of Technology (Breazeal 2000). It has auditory, visual and expressive systems intended to participate in human social interaction and to demonstrate simulated human emotions and appearance. It has not been used to perform any tasks apart from simulating social – and in particular emotional – interaction, using vocalization and facial expressions that are created through movements of the ears, eyebrows, eyelids, lips, jaw, and head.

Sage (also called *Chips*) concerns a series of interactive museum tour-guide robots that can demonstrate affective behavior, developed at Carnegie Mellon University (Burgard et al. 1998; Nourbakhsh et al. 1999). It has the ability to avoid obstacles and navigate robustly and has been functioning in several museums in the US and Europe. It can demonstrate moods in response to occurring events, like frustration when someone blocks the robot on its way though the museum. It can only express its moods and emotions vocally.

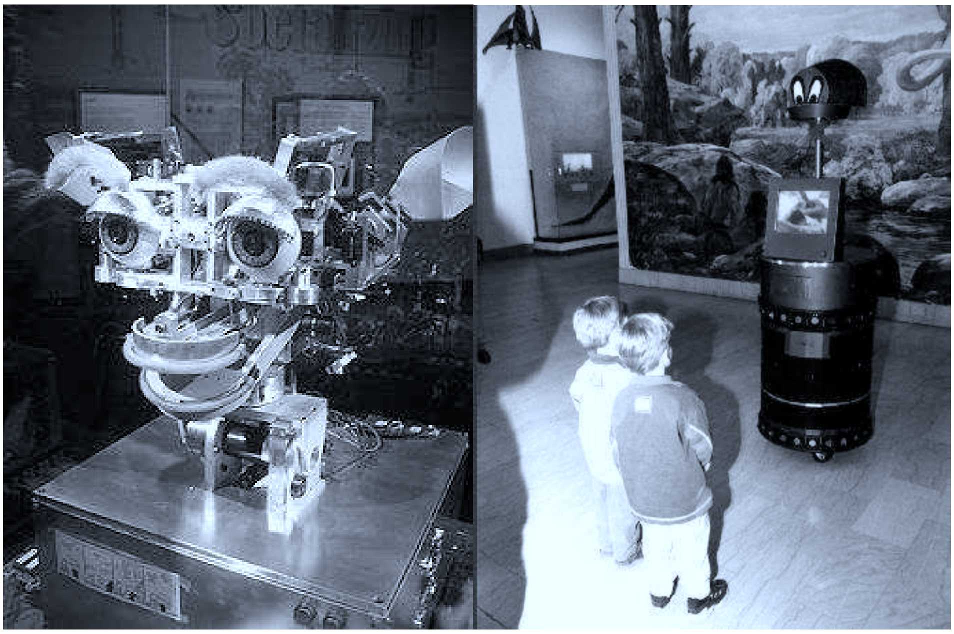


Figure 2.1. Kismet and Sage

As we have stated earlier, robots that have been developed to be assistive are not necessarily social robots. In fact, one of the earliest projects in the field of assistive robotics, MOVAID (Mobility and activity assistance systems for the disabled) (Dario 1999) is actually not a social robot (see Figure 2.2). The MOVAID robotic system consists of a number of fixed workstations in various locations in a home (such as the kitchen and the bedroom) as well as a mobile robotic unit able to navigate in the environment while avoiding unexpected obstacles, grasp and manipulate common objects, and dock to the fixed workstations to exchange its data and recharge its power supply. The robot has a camera and graphical interface to interact with the user, allowing him/her not only to monitor what the robot is doing, but also to collaborate with it by indicating objects and positions on the screen.

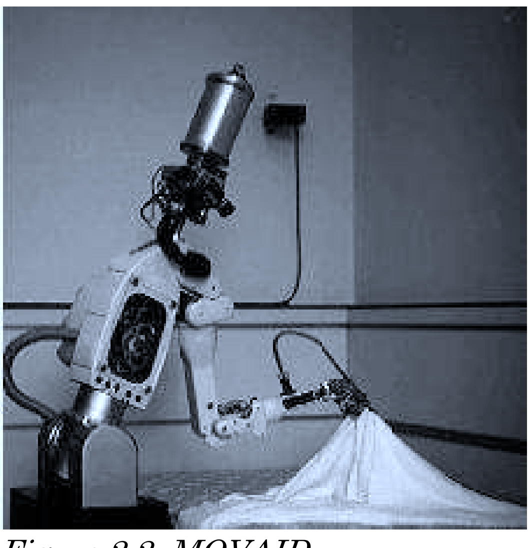


Figure 2.2. MOVAID

Robot research in eldercare features different types of assistive robots that we all are non-industrial robot types (industrial robots are focused on production processes).

As Figure 2.3 shows, assistive robots can be divided into two subcategories: they can either be non social robots or social robots. The first type concerns physical assistive technology that is for example developed for rehabilitation and that is not in any way socially interactive. Examples are the above mentioned MOVAID system, intelligent robotic wheelchairs (Gomi and Griffith 1998), artificial limbs and exoskeletons (Kazerooni 2005). The second type of assistive robots is socially interactive. These robots are systems that can be perceived as social entities that communicate with the user or are communicated with by the user as such (including touching and sensing). Of course there is an overlap between these two categories, since there are also projects on social robots that are to be used for rehabilitation purposes (Tapus et al. 2007), but generally in robotics these are separated fields.

Assistive social robots can also be divided in two subcategories. First, there are robots that we will refer to as *service* robots. They are used as functional devices and are not primarily designed for social support. Functionalities are related to the support of independent living by supporting basic activities (eating, bathing, toileting and getting dressed) and mobility (including navigation), providing household maintenance, monitoring of those who need continuous attention and maintaining safety. Examples of these are 'nursebot' Pearl (Pollack et al. 2002), the Dutch iCat (although not especially developed for eldercare) that we will describe in chapter 3 and the German Care-o-bot (Graf et al. 2004). Also categorized as such could be the robot used in the Italian RoboCare project, which is developed as part of an intelligent assistive environment for elderly people (Bahadori et al. 2003). A similar project is *Intelligent Sweet Home* at KAIST in Korea (Mukai et al. 2008; Park et al. 2008). They developed a 'steward robot' called *Joy* and a screen agent to be part of an intelligent living environment. All assistive social robots mentioned here, are described more extensively in Appendix A.

A second type of assistive social robots used in eldercare is what we will refer to as *companion* robots: they provide pet-like companionship which is possibly beneficial to the health and wellbeing of elderly users, but they do not provide functional assistance. Examples are the Japanese seal shaped robot Paro (Wada et al. 2003a), the Huggable (Stiehl et al. 2006) (both especially developed for experiments in eldercare) and Aibo (a robot dog by Sony, see Table 2.1).

These two types are not exclusive, for many robots can hardly be categorized strictly in either one of these two groups. Some companion robots (for example Aibo) can also be programmed to perform service type activities (Bartlett et al. 2003) and some service type robots (like both Pearl and iCat) can provide companionship. In addition, to be accepted by elderly users it could be helpful for an assistive robot to have some social abilities that would be typical for companion robots.

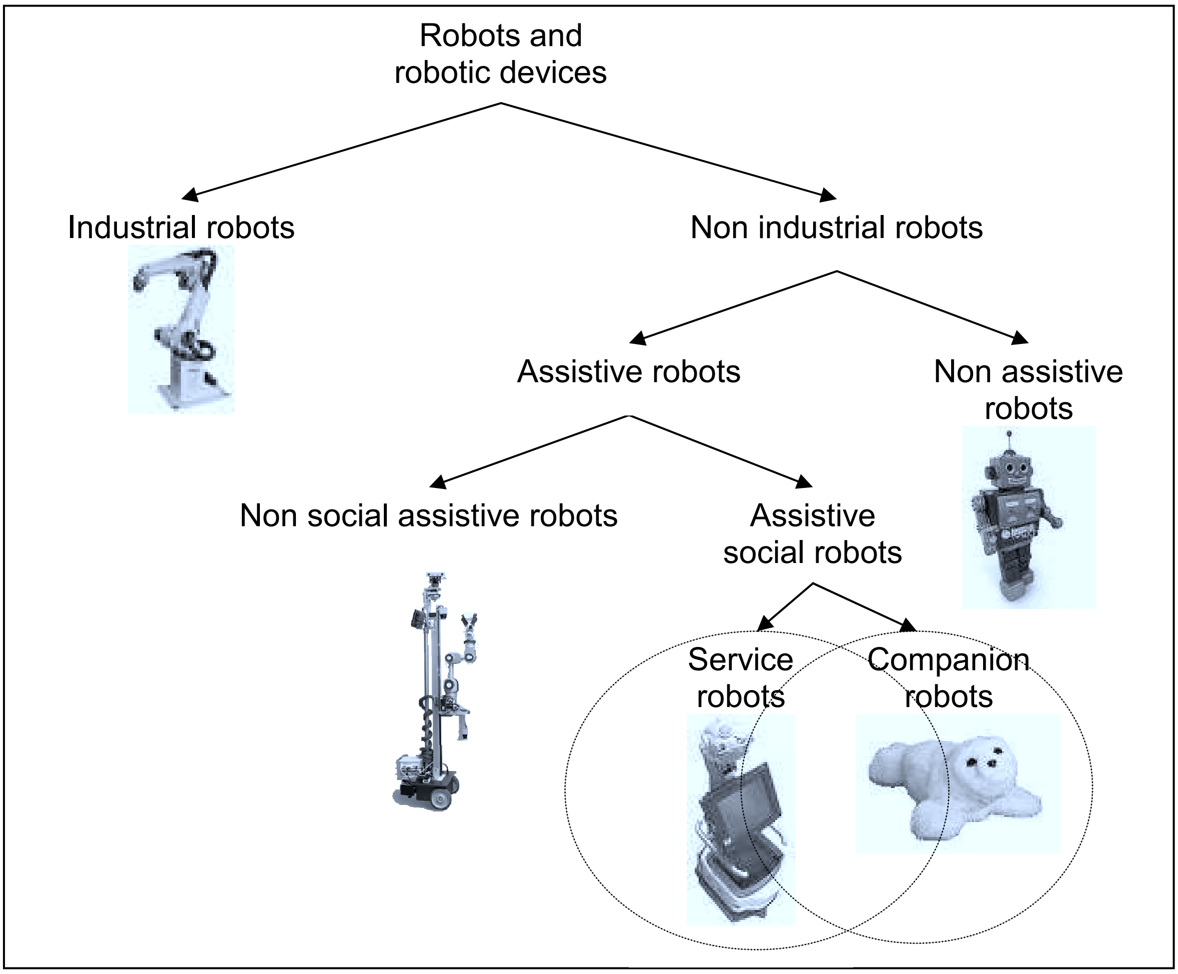
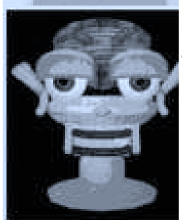
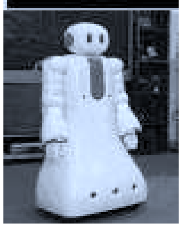


Figure 2.3. Categorization of robots

We mentioned earlier that screen agents also fit within our definition of assistive social robots. The screen agents that have been developed for elderly users can all be categorized as service type robots. Laura, for example, is an interface to a system (called Fittrack) that stimulates older adults to do physical exercises (Bickmore 2003) and Steffie is developed as a part of a website (www.steffie.nl) where she features as a talking guide for older adults, explaining the internet, e-mail, health insurance, cash dispensers and railway ticket machines. Also the earlier mentioned Korean Intelligent Sweet Home features a screen agent that functions as an interface to the automated home (Mukai et al. 2008; Park et al. 2008).

Table 2.1 provides an overview of robots and screen agents that have been developed for older adults or have been used for user studies concerning older adults. The list is up to date until 2008 and contains only those systems that have been described in literature, which means it is not exhaustive. It does however contain the systems that we have used in the experiments, described in this thesis (Annie, iCat, RoboCare and Steffie).

Robot	**Type**	**Embodiment**		**Interaction**	**Mobility**	**Functionalities**
Aibo	*Companion*		*Robot, dog*	*Sound Body movements*	*Walking*	*Companionship*
Annie	*Service*		*Screen agent, Female Humanoid.*	*Speech i/o Gestures Facial expressions*	*Not mobile*	*Monitoring Controlling devices Providing information*
Care-o-bot	*Service*		*Robot*	*Speech i/o Screen i/o*	*Wheels*	*Butler Guide Physical aid*
Homie	*Service*		*Robot, dog*	*Sound Body movements*	*Not mobile*	*Communication Companionship*
Huggable	*Companion*		*Robot, bear*	*Speech i/o Body movements*	*Not mobile*	*Companionship*
iCat	*Service*		*Robot, cat*	*Speech Facial expressions*	*Not mobile*	*Monitoring Controlling devices Providing information*
ISH – software agent	*Service*		*Screen agent*	*Speech i/o Facial expressions Screen i/o*	*Not mobile*	*Controlling devices Providing information*
ISH - Joy	*Service*		*Robot, part of intelligent home*	*Speech i/o Gestures*	*Wheels*	*Controlling devices Providing information Physical aid Butler*

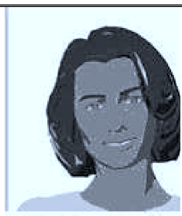
Laura	*Service*		*Screen agent, Female Humanoid.*	*Speech i/o Gestures Facial expressions*	*Not mobile*	*Coaching Providing information*
Nursebot	*Service*		*Robot (with touch screen)*	*Speech i/o Facial expressions Screen i/o*	*Wheels*	*Guide Providing information*
Paro	*Companion*		*Robot, seal*	*Sounds Body movements*	*Not mobile*	*Companionship*
Ri-man	*Service*		*Robot, android*	*Speech i/o Screen i/o*	*Wheels*	*Physical assistance*
Robocare	*Service*		*Robot, part of intelligent home*	*Speech i/o Screen output*	*Wheels*	*Guide Physical aid Butler*
Steffie	*Service*		*Screen agent, Female Humanoid.*	*Speech output Screen input Gestures Facial expressions*	*Not mobile*	*Providing information*

Table 2.1. Overview of assistive social robots for older adults (Appendix A provides a more extensive overview)

The table shows that there is a variety, including zoomorphic and anthropomorphic systems. Three of the four screen agents are female and all companion type robots are zoomorphic.

2.3 Overview of user studies on assistive social robots

We concluded in section 1.4 that there is very little quantitative research on acceptance of social robots by older adults. Still, as we have seen in the previous section, there are quite some systems that have been developed. Some of these systems have been subject to user studies that can give some information on the effect that social robots have on elderly users. To establish an overview of the results of these studies and establish how it relates to our project, we made an inventory of published studies on the effects on elderly users or the response by elderly users concerning the robots that have been presented in the previous section. This overview - which does not contain publications based on our own research - is presented in a summarized version in Table 2.2 and in an extended version in Appendix B.

We summarized the design of the study, the type of measurement and the result of the study. A positive result (+) can be a positive effect on health or wellbeing or an indication that the robot was positively received by the participants. If the symbol ± is used, this means there are both positive and negative effects. The design types (we found 3) and the varying outcome measures are listed by numbers, the legend can be found at the bottom of the table. The extended version in Appendix B contains extra remarks on design and outcome.

The table shows that most robots have been subject to user studies and most of those studies were case studies. Furthermore, it shows that the majority of user directed studies that have been carried out, featured Japanese toy robot dog Aibo and Paro, both zoomorphic robots and both companion type. This limits the possibilities to generalize the results, as it has been concluded that not only functionality, but also form and material does matter a lot for acceptance and effects of companion robots (see e.g., Kidd et al. 2006; Taggart et al. 2005).

A third conclusion is that most studies report positive effects of robots on participants, and the effects are diverse. Elderly react positively with respect to mood, health status, memory function and social connections with others. For example, companion robots seem to alleviate stress (e.g., measured by stress hormones in urine) and increase social interaction (measured by the frequency of contact between elderly). Nevertheless, these effect oriented studies concern companion type robots and are no direct indication of acceptance.

A fourth conclusion is that narrative records, present in a large portion of these studies, show that most elderly actually report to like the companion type robots (or their controls for that matter). A wide variety of research designs has been used, and these studies indicate a positive effect of companion robots on elderly, varying from expressed appreciation to improvement of health or wellbeing. However, it is not exactly clear what the cause of this effect is and whether this effect also would be applicable to service type robots. It would certainly be worth considering to add companion type features when designing service type robots and establish the influence of this on their acceptance.

A fifth conclusion is that 24 out of 31 found studies are done in Japan (mostly by the same group of researchers). It has been shown that robot perception is culturally dependent in a study comparing the measured attitude for participants different nationalities (Bartneck et al. 2007) where the attitude of Japanese participants differed in many aspects from the attitude of participants from other countries (e.g. US, Mexico). The Japanese were not as positive as stereotypically assumed: results indicated that the Japanese are concerned about the impact that robots might have on society and that they are particularly concerned with the emotional aspects of interacting with robots. The results of the 24 Japanese studies should therefore not be generalized too easily to other cultures.

Reference	Design	N	Result	Measures	Term of use
Aibo					
Kanamori et al. 2002	2,4	3	+	1,5	?
Kanamori et al. 2003	2,4	5	+	1,3,4,5	7 weeks
Mival et al. 2004	3,4	10,12	+	6	
Suga et al. 2002	2	23	+	1	2 months
Sakairi 2004	2,4	8	+	3,5	30 minutes
Suga et al. 2003	2,4	15	+	1	?
Tamura et al. 2004	2,4	?	±	3	5 min. intervention
Turkle et al. 2006	2,4	2	+	5,6	several months
Yanagi & Tomura 2002	2	46	+	5	Several hours
Care-o-bot					
Graf et al. 2004	2,4	6	+	5	
iCat					
Looije et al. 2006	3,4	6	±	5,6	< 1 hour
Homie					
Kriglstein and Wallner 2005	2,4	2	+	3, 6	?
Laura/Fit track					
Bickmore & Picard 2005	3,4	8	+	6,7	?
Paro					
Giusti and Marti 2006	2,4	5	±	3	1 month
Kazuyoshi et al. 2003	2,4	12,11	±	2	3 weeks
Kidd et al. 2006	1,4	23	+	3,2	4 months
Marti et al. 2006	2,4	1	+	3	1 time
Saito et al. 2002	2	20	+	1	6 weeks
Saito et al. 2003	2,4	12,11	-	5	3 weeks
Taggart et al. 2005	1,4	18	+	3	20 minutes
Wada et al. 2002a,b	2	11	±	2	3 weeks
Wada et al. 2003a-d	2,4	3-12	+	1,2	3 weeks,
Wada et al. 2004a-c	2	10,11,12	+	1	3-14 weeks
Wada et al. 2005a,b	2	23	+	2,3	1 year
Wada et al. 2005c	2,4	8	+	2	17 months
Wada et al. 2005d	2	14	+	5	20 minutes
Wada et al. 2006	2	14	+	2	10 weeks
Wada & Shibata 2006,2007	2	11,12	+	1,3	1 month
Pearl					
Montemerlo et al. 2002	2,4	6	+	5	5 days
Pineau et al. 2003	2,4	6	+	5	5 days
Robocare					
Giuliani et al. 2005	1	123	±	6	?

Design	**Outcome measures**	
1. Comparative cohort	1. Health status	5. Other design criteria
2. Case studies	2. Mood	6. Remembering
3. Focus group	3. Communication	7. Acceptance rating
4. Narrative records	4. Loneliness	

Table 2.2. Summarized overview of user studies on assistive social robots

A sixth conclusion is that most studies have very low sample sizes, which limits the possibilities for statistical processing to provide reliable answers to research questions.

Our last conclusion in relation to this overview is that the studies are in many cases not conclusive in their outcome. Sometimes the results are difficult to interpret in the sense that the control condition (e.g., a fake Paro) has an effect that is more or less the same as the effect in the experimental condition, or the number of participants is too small to have significant findings based on quantitative data (Saito et al. 2003; Wada et al. 2003a; Wada et al. 2003b; Wada et al. 2003c; Wada et al. 2003d; Tamura et al. 2004). Moreover, some studies are contradictory in terms of their outcome (Saito et al. 2003; Taggart et al. 2005).

Based on the found literature we may conclude that there is some evidence that companion type robots have positive effects in healthcare for elderly with respect to mood, health status, memory function and social connections with others. However, the publications we found on quantitative studies often had very low sample sizes and mostly did not assess robot acceptance. The only study addressing acceptance of a social agent by older adults that included quantitative analysis is a study on screen agent Laura, featuring only 8 participants (Bickmore et al. 2005). We will discuss acceptance methodology and user studies on acceptance of robots further in the next sections.

2.4 User acceptance studies

Research on why individuals adopt (or do not adopt) new technologies has lead to several types of research that can be divided into three main streams based on their focus: implementation oriented, design oriented and perception oriented.

The first stream has focused on measuring the success of the implementation of new technology within an organization (Yoon et al. 1995; Gallivan 2001; Frambach and Schillewaert 2002). An example is the research by Leonard-Barton and Deschamps which identifies different employee characteristics (like personal innovativeness or task related skills) rather than perceptual factors and establishes how these different characteristics mediate the influence of implementation management strategies (Leonard-Barton and Deschamps 1988). Another example is the research by Orlikowsky, which focuses on the process of technology implementation, shaping and being shaped by human actions and interacting with organizational structures (Orlikowski 1992; Orlikowski 2000).

A second, design oriented, stream focuses on usability linked task-technology fit which is common in the field of HCI (Goodhue 1995; Goodhue and Thompson 1995). It defines the tasks that the user wishes to perform and evaluates users' responses during or after usage of new technology (often a prototype) from a developer's perspective. The results can be used to establish how well (how fast, how easy, how efficiently) a user can use the technology for the tasks and define the focus for improvements.

The third stream is oriented towards user perception (Davis et al. 1989; Compeau and Higgins 1995). It aims to identify the influences on individual acceptance, which is measured by quantifying the expressed *Intention to Use* a specific technology by test participants. This Intention to Use is both used as dependent variable which is determined by established influences and as a predictive influence (determinant) on the actual use of the system. The influences predicting Intention to Use, form a cognitive model that can be validated by establishing that it indeed predicts actual use. An identified influence could for example be the *Perceived Usefulness* of a device: the higher this influence is, the more likely the technology will be accepted. Statistical procedures can be used either to establish the strength (significance) of a hypothesized influence or to establish the dominant influences on acceptance of a certain technology by a specific user group.

Although our findings can be of use from a managers or developers perspective, we are interested in the perceptual processes of acceptance of technology in a non working environment. We therefore find that the first and second stream approaches are either restricted to a working environment or to a strictly task based technology, which makes the third stream more favorable for our present research in which we want to identify the influences on acceptance of robots performing very diverse tasks, and by a specific group in a non-working environment. This means we want to understand usage as a dependent variable and will connect to the third stream, often referred to as ‘technology acceptance modeling’.

In the next sections, we will describe this research stream in more detail by explaining its basic assumptions and its different models. Subsequently we will discuss its application to robotic systems.

2.4.1 Basics of technology acceptance modeling

Technology Acceptance Modeling (TAM) has had much attention in the last two decades. Although there have been earlier models, an overview of technology acceptance usually starts with the introduction of the technology acceptance model (TAM) by Davis in 1986 (Davis 1986; Davis 1989). This model has become one of the most widely used theoretical models in technology evaluation. It was adapted from the Theory of reasoned action (TRA) by Fishbein and Ajzen (Fishbein and Ajzen 1975; Fishbein 1980).

As visualized in Figure 2.4, TRA states that a person’s intention to perform a certain behavior (Behavioral Intention) is defined by two influences: Attitude (an individual’s positive or negative feeling about performing the target behavior) and Subjective Norm (the person’s perception that most people who are important to him think he should or should not perform the behavior in question).

It has been applied in very different studies, including dieting (Sejwacz et al. 1980), using condoms (Greene et al. 1997) and consuming genetically engineered foods (Sparks et al. 1995).

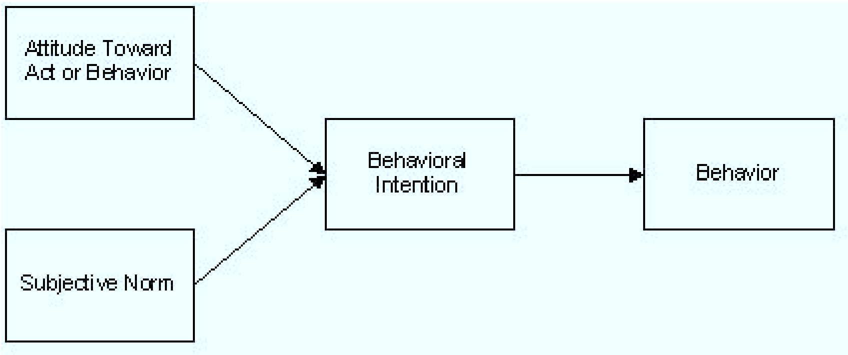


Figure 2.4. Theory of reasoned action assumptions (source: Fishbein and Ajzen 1975)

To apply this model to acceptance of information technology, Davis defined attitude as 'the accumulation of feelings about the usefulness and the ease of use of the specific technology' and initially dropped the notion of Subjective Norm. This means that, given this choice, the model in its most basic form states that Perceived Usefulness and Perceived Ease of Use are influences that determine the (Behavioral) Intention to Use a system and it poses that this outspoken intention is predicting the actual use (see Figure 2.5).

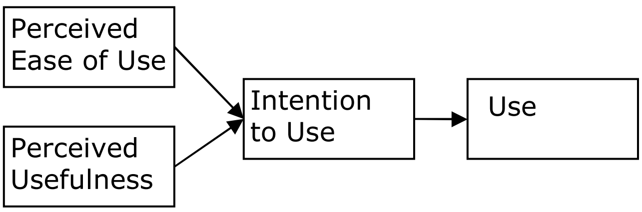


Figure 2.5. Basic TAM assumptions

Usually in acceptance model methodology, each influence is represented in a questionnaire by a group of items (questions or statements) which can be replied to on a five or seven point (occasionally six, eight, nine or ten point) Likert type scale. To the answers on these questionnaire, scores can be attributed so that statistic processing is possible. The influences, thus represented by variables derived from questionnaire scores, are usually called constructs.

2.4.2 Further development of TAM

The original TAM model as it had been developed by Davis has been used for many different types of technology and has been extended with more influences that were found to influence Intention to Use or Usage. In a later stage, Davis and Venkatesh reintegrated the concept of Subjective Norm. This model is usually referred to as TAM-2 (Venkatesh and Davis 2000).

Since the 1990's other influences have been explored and different interpretations of influences that were recognized in the Theory of Reasoned Action and TAM have been studied. The Motivational Model, for example, focused on Intrinsic Motivation and Extrinsic Motivation as determinants (Vallerand 1997) and the Theory of Planned Behavior (Ajzen 1991; Taylor and Todd 1995) extended the TRA-model with Perceived Control (perceptions of internal and external constraints on behavior). In the same period, the Model of PC Utilization (Thompson et al. 1991) introduced many alternative influences, among which Social Factors (peer group influence) and Facilitating Conditions (objective factors in the environment that facilitate use of the technology). Furthermore, the Innovation Diffusion Theory (Rogers 1995) has been applied to technology acceptance, introducing influences like Voluntariness of Use and Image (Moore and Benbasat 1991).

This summary of alternative models shows that there are many possible approaches to measure technology acceptance. It suggests that there are different angles to measuring acceptance and the most effective approach may also be dependent on the type of technology in question, the user group and the specific context.

2.4.3 UTAUT

In an attempt to construct a unifying model that incorporates the most widely used approaches, Venkatesh et al. (2003) evaluated eight theoretical models that employ intention and/or usage as the key dependent variable: the Theory of Reasoned Action, Motivational Model, Theory of Planned Behavior (TPB), a combined TAM and TPB model, Model of PC Utilization, Innovation Diffusion Theory, and Social Cognition Theory. All these theoretical models were empirically tested and compared. A key issue within that process was the predictive power of the model. This was measured by establishing how effectively the used influences determined the variable representing the Intention to Use the system. Next, a new model was developed, using the essential elements of these eight models and combining them if possible. Finally, this new model was empirically validated in a user study in which both Intention to Use and actual use were measured within the context of information technology acceptance in a working environment. The explanatory power of the new model was much stronger: while the contributing models would explain between 17 and 53 percent of the variance in Intention to Use the system, the new model was found to explain up to 70 percent of the variance in Intention to Use.

The result of this process is the UTAUT (Unified Theory of Acceptance and Use of Technology) model which has also been used in previous research in acceptance of robots (De Ruyter et al. 2005; Looije et al. 2006). It states the following influences to be direct determinants of *Intention to Use* (also called *Behavioral Intention*; see Figure 2.6):

- Performance expectancy, a construct similar to Perceived Usefulness, but defined in a broader sense as 'the degree to which an individual believes that using the system will help him or her to attain gains in job performance'.
- Effort Expectancy, a construct similar to Perceived Ease of Use, but also more broadly defined as 'the degree of ease associated with the use of the system'.
- Social Influence, defined as the degree to which an individual perceives that important others believe he or she should use the new system.

Actual use of the technology is determined by Intention to Use and Facilitating Conditions. The latter is defined as the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system.

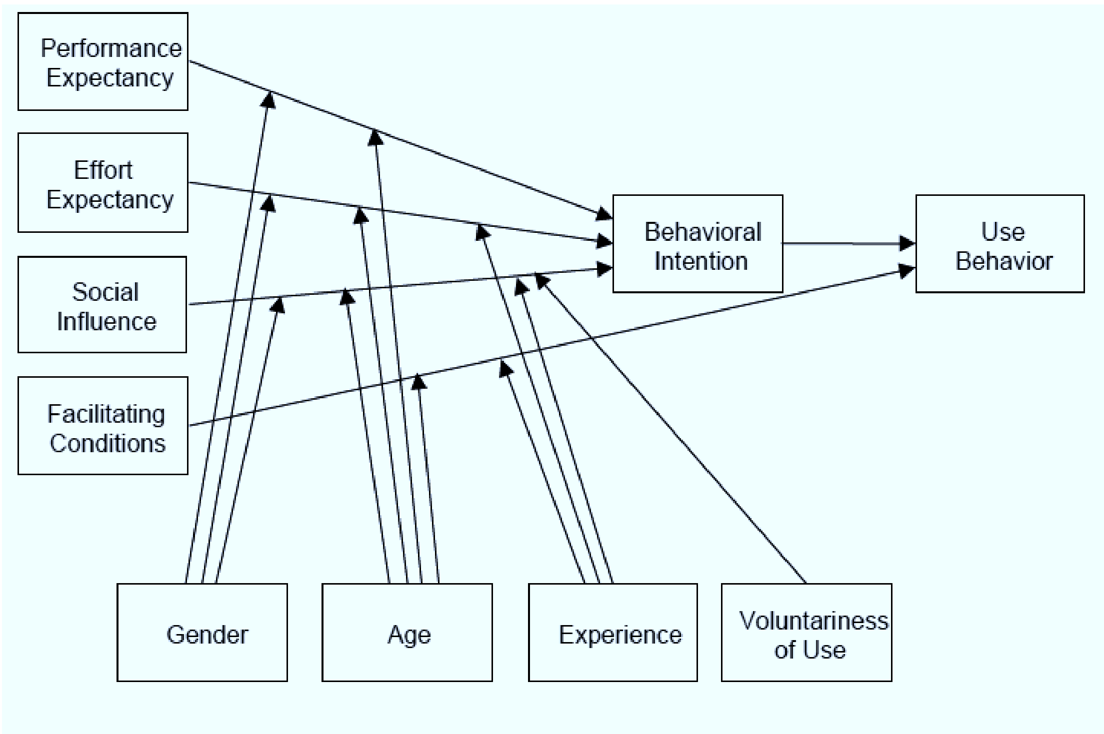


Figure 2.6. UTAUT model: direct influences and moderating factors – indirect influences (Attitude, Anxiety and Self-efficacy) are not represented.
(source: Venkatesh et al. 2003)

Besides these constructs, the model features four moderating influences that make the effect of a construct on Intention to Use stronger or weaker: Gender, Age, Experience and Voluntariness of Use. Their influence is established as follows:

- The influence of Performance Expectancy on (Behavioral) Intention to Use is moderated by Gender and Age, such that the effect will be stronger for men and particularly for younger men.

- The influence of Effort Expectancy on Intention to Use is moderated by Gender, Age, and Experience, such that the effect will be stronger for women, particularly younger women, and particularly at early stages of Experience.
- The influence of Social Influence on Intention to Use is moderated by Gender, Age, Voluntariness, and Experience, such that the effect will be stronger for women, particularly older women, particularly in mandatory settings in the early stages of Experience.
- The influence of Facilitating Conditions on Usage is moderated by Age and Experience, such that the effect will be stronger for older workers. However, it decreases with increasing Experience.

Besides the four influences on Intention to Use, three other constructs have been studied by Venkatesh et al.

- Attitude towards using technology, defined as an individual's overall affective reaction to using a system.
- Self-efficacy, defined as the judgment of one's ability to use a technology (e.g., computer) to accomplish a particular job or task.
- Anxiety, defined as Evoking anxious or emotional reactions when it comes to performing a behavior (e.g., using a computer).

These constructs have however been established to be no direct determinants of Intention to Use, for their influence is 'captured' by the stronger influences of the above constructs. It is however possible that they have an indirect influence within the model: they may be determinants of the four constructs that determine Intention to Use (the "direct influences"). However, although this possibility has been recognized it is not further explored by Venkatesh et al. (these "indirect influences" are also not represented in the visualization of the model in Figure 2.6). As we will see in chapter 4, more recent studies led to a more prominent position of the influence of attitude as it is often found to be a determinant of Intention to Use and sometimes as a determinant of actual use (Yang and Yoo 2004; Knutsen 2005; Pynoo et al. 2007).

The questionnaire (statement list) as used with the UTAUT model is presented in Appendix D. It is indicated by Venkatesh et al. that items in this list can be slightly modified when adapting to a specific technology or user group and this has been done on several occasions (Li and Kishore 2006; Hennington and Janz 2007; Wu et al. 2007). However, when larger changes are made or even constructs are added a new validation would be required. The validation of a model typically includes a long term observation of the actual use of the technology, which makes it possible to relate scores on Intention to Use to Usage of the system and establish whether the latter is indeed predicted.

After the UTAUT model has been published, it has not only been applied to different types of technology, but also has been criticized and tested with alterations. Many studies conclude that the model cannot be applied universally, but needs extensions or modifications when applied to technologies or target groups that are different from the context (acceptance of information technology in a working environment) in which it has been initially developed and validated. Some examples:

- Studies on acceptance of e-business related systems (Cody-Allen and Kishore 2006) suggest that this particular context justifies the extension of the model with constructs concerning Perceived Quality, Trust, and Satisfaction (Knutsen 2005; Uzoka 2008).
- A study concerning hybrid media applications, in particular bar code reading applications for camera phones (Louho et al. 2006) demonstrated Attitude to be a strong direct influence on Intention to Use for a group of 20 users. This determining influence of Attitude is confirmed by many other studies (e.g. Yang and Yoo 2004; Knutsen 2005; (Pynoo et al. 2007)).
- A study by Schaper and Pevan (Schaper and Pervan 2007) on ICT-acceptance by Australian health professionals adds not only Attitude, but also a newly developed construct called Compatibility. It shows that Intention to Use in this context is more accurately predicted when these constructs are added as determining influences. Moreover, they demonstrate that a better understanding of acceptance can be achieved when also indirect influences (determinants of the direct influences) are included. In their model these are Anxiety, Self-efficacy and Compatibility (which is thus both a direct and indirect determinant).

These examples illustrate that although Venkatesh et al. present the UTAUT model as complete and universally applicable, some studies suggest that it needs to be adapted to be applicable in a context that is different from the one it has been developed in. Moreover, many studies suggest that the construct of Attitude should have a more prominent position.

2.4.4 Statistical techniques

As we explained, in technology acceptance modeling, each influence represented in a model is a variable. The values for these variables, including the variable representing Intention to Use, are mean scores on questionnaire items. After these values are calculated, several statistical methods can be used. In general these concern the following goals:

- Test the reliability of the constructs. It has to be established that questionnaire items that are joined within a construct have a similar pattern in the scores.
- Confirm hypothesized relationships between constructs. Hypotheses can be based on theoretical considerations or on findings in related research. They can concern direct determinants of Intention to Use, or indirect influences determining these direct determinants. (Sun and Zhang 2006)
- Establish the influential strength of different constructs. If there are several determinants of Intention to Use within a model, it is important to establish which of these have a stronger influence. The outcome of this can give some specific information related to the technology or the user group of a specific study.

- Explore alternative relationships between the constructs of a model. Constructs may very well have a determining influence on each other that has not been hypothesized before the analysis.
- Establish the significance of moderating influences. The significance of these influences vary for each context, technology and user group (Sun and Zhang 2006) so if applicable (they are not always subject of research in TAM development projects), this has to be established for each set of data.

The most commonly used statistical techniques are the following:
- To test the reliability of the constructs, Cronbach's Alpha is computed (Santos 1999): scores on the items of a construct are processed to see whether their patterns are sufficiently similar. A reliable grouping of items would have an alpha of at least .7 (Nunnaly and Bernstein 1978). Sometimes items are deleted to obtain a score that passes this threshold. Furthermore, a principal component analysis with rotation component matrix can be used to check whether items that belong to a construct indeed 'load' on the same factor (see section 4.3).
- To confirm or reject hypothesized dependencies between constructs and to establish the influential strength of different constructs, linear regression analysis can be used (Montgomery et al. 2001; Tabachnick and Fidell 2001). It can be used to confirm or reject theoretical relationships between one dependent variable (construct) and one (simple linear regression) or more (multiple linear regression) determining influences (constructs). If more influences are tested, it indicates how strong each influence determines the dependent construct. A linear regression analysis demands preferably at least 20 participants for each construct, but never less than 5. It also results in a coefficient of determination, the R^2 value (between 0 and 1), which indicates how strongly the variance within the dependent variable is explained by all the influencing variables. Within an acceptance model, the R^2 value of the influences on Intention to Use is used as an indication of the reliability of the model.
- The most profound way to analyze results would be to apply structural equation modeling which tests all relations in one analysis. This can be used both to establish alternative paths (explorative) and the strength of presumed construct interrelations (confirmative). This would, however, demand at least 15 to 20 cases (users) per construct, and in this field it is often not possible to meet the challenge of gathering this number of participants (see section 8.4).
- To explore alternative relationships between constructs or items, sometimes a correlation analysis (Spearman or Pearson) is carried out, although correlation does not indicate causality. It can be done with any number of participants. Correlation analysis is especially useful if multiple tests need to be compared
- Of course additional statistics can be computed. When comparing different conditions or user groups for example, a T-test or Mann-Whitney U-test can be carried out, besides ANOVA. An example of a statistical technique that is used if more than two groups are compared is a post hoc Games and Howell (see section 6.4).

- To establish the significance of moderating factors UNIANOVA (Zhu and Mao 2008) or a Chow's test (Chow 1960) are used. Both tests have similar procedures. If construct A is found to determine construct B, they establish whether this determining process is significantly moderated by factor C.

2.5 Acceptance studies concerning robots

Since technology acceptance methodology has mainly been developed to be applied in a working environment, it is not surprising that it is hard to find examples of acceptance studies in which assistive robots and screen agents are involved: this type of technology has not been applied frequently to a working environment. In addition, evaluation of this technology is usually focused on functional aspects of the system itself and how well users can work with it (Clarkson and Arkin 2007; Feil-Seifer et al. 2007), while acceptance methodology focuses on how users perceive it and on establishing influences that determine acceptance.

Nevertheless, some examples can be found. For instance, robot interaction research in which acceptance did play a significant role is described by De Ruyter et al. (De Ruyter et al. 2005). It concerned the iCat (see paragraph 2.2) which was tested in a Wizard of Oz experiment where the robot was controlled remotely by an experimenter while it was suggested that the robot was autonomous. This experiment was done in a laboratory setting, with adult, but not elderly participants.

The participants were asked to program a DVD-recorder and to participate in an online auction, by using the iCat interface. They were exposed to an introvert and an extravert version of the iCat interface to see whether this difference in interaction would lead to different scores in degree of acceptance. To measure acceptance, the UTAUT questionnaire was used. The aim of the study was to find out to what extent participants would use the iCat at home after having experienced it. To see whether participants would perceive the extravert iCat to be more socially intelligent, a social behavior questionnaire (SBQ) was developed and used. The results showed that the extravert iCat was indeed perceived to be more socially intelligent and that this version also was more likely to be accepted by the user.

The same robot was used in an experiment by Looije et al. (2006) where it featured as a personal assistant for a small (n=6) group of people with diabetes. It was compared to a text interface and the robot was used in a more and a less socially intelligent condition. Results showed that participants preferred the robot over the text interface and appreciated a more socially intelligent robot over a less socially intelligent one.

An example of research involving screen agents for elderly users is reported by Bickmore et al. (2005). It concerns the evaluation of elderly users' responses to relational agent (a screen agent that simulates to have a personal interest in the user) called Laura, which functioned as a health advisor for older adults. This screen agent's interaction consisted entirely of a relationship-building dialogue including social dialogue, meta-relational dialogue, empathy exchanges, humor, and reciprocal self-disclosure exchanges. The studies' findings indicate that the agent was accepted by the participants as a conversational partner on health and health behavior. It was seen as trustworthy and friendly and it was also found to be a good health advisor. Other research with the same agent by Bickmore and Schulman (Bickmore and Schulman 2006; Bickmore and Schulman 2007) focused on the screen agent's ability to participate in long term relationships. This ability is linked to the notion of social presence (Lombard and Ditton 1997; Lee and Nass 2003) that people feel in interaction with systems, which can play a role in interpreting the responses of people when social abilities are perceived. It was found that even though subjects conducted an extended dialogue with the agent, they apparently fail to develop a strong social bond. These studies carefully suggest that in order to achieve acceptance of a social agent, social behavior is important but establishing a long term social relationship may not be needed.

2.6 Social acceptance, credible skills and conversational behavior

In the beginning of this chapter we addressed the sociability of assistive robots. We discussed the seven point list of social robot characteristics by Fong et al. (2003) and described social robots that were assistive. In our discussion of technology acceptance, however, the issue of sociability was not addressed, which is not surprising if we take into account that technology acceptance methodology has not been developed for socially interactive systems. Still, there are a few issues to address within the framework of our present study if we want to apply this methodology to this type of system.

First there is the issue of 'social acceptance'. The research by Bickmore et al. as described in the previous section demonstrates that social interaction can only be successful if conversational partners accept each other as such. This means it is important to not only address acceptance in terms of the functionality and technological features, but also in terms of relational or social aspects of the interaction. In other words, we should address social acceptance complementary to functional technology acceptance. In this regard we view social acceptance as a user feeling comfortable with a robot as a conversational partner, apparently finding its social skills credible and accepting social interaction as a way to communicate with the robot.

This credibility of social skills is addressed by several projects that concern the development of one or more characteristics of the seven point list of Fong et al. (see section 2.2). The first item on this list for example (express/perceive emotions), is addressed by numerous projects that focus on the affective abilities of robots. These abilities concern expression (Moshkina and Arkin 2005; Leite 2007) and perception of emotions by robots (Breazeal and Aryananda 2002; Yilmazyildiz et al. 2006; Kulic and Croft 2007), or both (Scheutz et al. 2005). These studies address emotion perception and expression by the development of theoretical frameworks or by experimenting with robots and screen agents in which these affective computing abilities have been implemented.

Sometimes a study on emotion perception and expression is combined with the development of abilities to use natural cues like gaze and gestures (Yoshikawa et al. 2006; Satake et al. 2009) and other social competencies, including distance keeping (Pacchierotti et al. 2006; Gockley et al. 2007). Also the development of dialogue related abilities is subject to recent projects that concern both theory and application (Foster et al. 2008; Kruijff et al. 2009). Furthermore, there are several projects concerning the study of human-robot relationship development (Jacobsson et al. 2008; Wagner 2009) and the impact of a robots personality on this (Konstantopoulos et al. 2008; Walters et al. 2008). These studies generally show that more social abilities lead to a more strongly felt positive relationship between user and robot and a higher perception of the usefulness of the robot.

Another issue that relates to both acceptance and the sociability of robots is the behavior of users during interaction with it. If a robot has credible social skills, this can have an impact on the user's behavior. It can be expected that the more a user accepts a robot as a conversational partner, the more he will show conversational engagement by demonstrating more expressive behavior – as it has been established that engagement and expressiveness interrelate (Cappella 1983; Coker and Burgoon 1987; Cacioppo et al. 1992; Xaverius and Mathews 2004). Based on this assumption, we can state that expressiveness is an indicator of conversational engagement, which in its turn indicates acceptance of a conversational partner. This would be a novel approach within the field of Human-robot interaction: although the relation between behavioral cues and engagement when interacting with computer systems has been explored by Axelrod et al. (Axelrod and Hone 2005; Axelrod and Hone 2006), using behavioral clues as an indication of acceptance is unprecedented.

2.7 Conclusions

From the overview of (experimental) robots used in eldercare (section 2.2), we concluded that there is some evidence for positive effects on health and wellbeing. However, it is limited in multiple aspects: most studies are done in Japan and considered a zoomorphic companion type robot and we often cannot be conclusive about the reliability of statistical results. In addition, there is no relation between these positive effects and acceptance.

Furthermore, we found that there is a unified model on technology acceptance (UTAUT) that has been applied to robot technology in a limited study, but not for elderly users. We also found that many studies show how the UTAUT model needed adaptation in the sense of extension, modification or both and since our context of assistive robots for elderly users is quite different from the one that has been used to validate UTAUT, it is likely we also need an adapted model. However, to be conclusive on this, we need to test the model first and establish its predictive power in the specific context of robots used by older adults.

A third conclusion is that we must take the notion of social acceptance into consideration as a concept that complements technology acceptance. This implies that research on robot and agent acceptance can be subdivided into two areas: (1) the acceptance of the robot in terms of usefulness and ease of use (functional acceptance) (Montemerlo et al. 2002; Pineau et al. 2003; Forlizzi et al. 2004; De Ruyter et al. 2005; Looije et al. 2006) and (2) the acceptance of the robot as a conversational partner with which a human or pet like relationship is possible (social acceptance) (Bickmore 2005; Wada and Shibata 2007). The experiments with companion type robots (like Paro and Aibo) were more focused on social acceptance while the experiments with service type robots (like Pearl and iCat) focused more on the acceptance of the robot regarding its functionalities.

To be able to obtain a complete view on acceptance of an assistive robot, we need a model that enables us to explore both social and functional acceptance. This means we also have to evaluate UTAUT on the accuracy in predicting both.

3. Measuring the influence of social abilities I

Parts of this chapter have been published earlier in (Heerink et al. 2006a; Heerink et al. 2006b; Heerink et al. 2006c; Heerink et al. 2009a)

3.1 Introduction

In the previous chapter we introduced the UTAUT technology acceptance model. We explained (in section 2.4.3) how the model has been developed by unifying existing models and incorporating all known influences. In the validation experiments by Venkatesh et al. (2003), it predicted the variance in the Intention to Use new technology in the workplace from 50% up to 70%. However, there has not been any evaluation of the model within a context of systems that can be perceived as socially interactive, nor has it been applied to elderly users.

In this chapter we will describe the setup and outcome of two experiments to evaluate the model, involving assistive social robots used by older adults. The adaptations of the UTAUT methodology concerned the specific context of older adults in a non-working environment and the specific systems (see section 3.6.3). With the outcome of these experiments we want to establish the robustness and explanatory power of the model and evaluate its potential to explore both social and functional acceptance.

For these experiments we used two different systems, a robot (iCat) and a screen agent (Annie), so we could establish whether the model would perform equally well for both. Moreover, we assumed that if we would have similar conclusions for both systems, this would be a good indication that our findings could be generalized for assistive agents.

Furthermore, we created two conditions for each system: a more social and less social version. This also enabled us to establish whether the model would perform equally well in these different conditions. If this is not the case, this could indicate that the model is incomplete.

Using these different conditions also enables us to evaluate an assumption mad in the previous chapter. We concluded that social acceptance and social abilities are related: a more social robot can be expected to be accepted better. If this assumption is correct, this should be reflected in the scores for Intention to Use.

As a manipulation check, we used control questions which formed the additional construct of Perceived Social Abilities. Analyzing the responses, we could establish whether the different conditions were indeed perceived as more and less social. Also this enabled us to establish whether the model would be able to explain eventual differences in the ways these conditions were perceived.

Furthermore, we added behavior observation to establish whether the difference of the conditions was reflected in differences in user behavior.

In sections 3.2 to 3.4 and 3.6 we will present the core concepts and assumptions, define our hypotheses and describe the used systems and experimental methods. In section 3.5 we will describe a pilot experiment and the lessons we learned from it and in section 3.7 we will present the experimental results. In sections 3.8 and 3.9 we will evaluate these results and present our conclusions.

3.2 Social characteristics, social abilities and acceptance

In our first chapter, we defined social robots by using the description of Fong et al. (Fong et al. 2003) of socially interactive robots. In the second chapter, we gave characteristics of social interaction that apply to many different types of robots that interact in very different ways. This means that different robots may be defined as social robots although they have very different abilities. The robot Paro, for example has a limited set of interactive possibilities (a purring sound, eye movements and body movements) that we can still categorize as social interaction. Robots like iCat and Pearl on the other hand, can not only produce sounds and move their eyes, but are also able to listen and speak, have different types of facial expressions (especially iCat). Thus, although they share the category of social robots with Paro, they have a different (and larger) set of social abilities (in section 3.4.3 we will further specify the possible differences in implemented social abilities for robots and screen agents).

When programming these robots with more advanced social abilities, these abilities can be manipulated. This means we can not only define a social robot as a robot that has some social abilities, but also classify robots that are more and less advanced in their abilities and we can increase or decrease the social abilities for a specific robot to study the effect of this manipulation.

There is, however, a potential discrepancy between *implemented* social abilities and *perceived* social abilities. The first type concerns the programmed behavior: we can give a robot the ability to show expressive behavior, to form polite sentences, etcetera. *Perceived* social behavior however, can be any type of behavior that is interpreted as social by the human interacting with the robot even if it is not intended to be social by the robot or its programmer (Moshkina and Arkin 2005). In addition, behavior that is intended to be social may not be perceived as such.

Studies show that in human-human interaction a higher rated set of social abilities correlates with a higher score on constructs related to social acceptance. Studies described by Giles and Powesland (1975) for example, found sophisticated social competencies (e.g. using the right tone) to be related to a positive social evaluation of subjects and studies described by Haslett (1990) found a relation between social competencies and attributed social status. Other studies demonstrate the relation between social acceptance and social skills for young children and scholars (Harter and Pike 1984; Vaughn et al. 1990; Evans 1992).

This relation between social abilities and acceptance could be similar for human-robot interaction: a higher developed set of social abilities would lead to a higher degree of acceptance – although it still has to be established whether this concerns social acceptance, functional acceptance or both (see section 2.6).

The importance of (perceived) social abilities, especially in a health- and eldercare environment, is also stressed by recent studies that focus on interaction with robots. A study by Cappella and Pelachaud (2001) for example, shows the positive effect of affective responsiveness of robots in social interaction on their appreciation, and a study by Breazeal (2003) demonstrates the positive effect of conversational skills (e.g. turn taking and expressive feedback) on the human-robot relationship - a topic that is further explored by Kidd and Breazeal (2005). The role of sociability in robots in particular when used for older adults is discussed by Forlizzi (2005) who concludes it to be essential, based on outcomes of related studies and theoretical considerations. Furthermore, it is explored (although with a Paro robot which has limited social skills) by Kidd et al. (2006) and Wada and Shibata (2007).

What these studies have in common is the finding that a robot will be more effective in its communication if it has a more advanced set of social abilities. This means that it can be expected to be easier and more pleasant to interact with and therefore would indeed be accepted more easily. Our present study with a more and less sociable condition could help us establish whether this assumption is right by showing a difference in acceptance that can be related to a difference in social abilities - or it could demonstrate that social abilities are of no influence on acceptance.

Moreover, we expect the study to establish whether the UTAUT methodology is able to predict acceptance equally well in conditions concerning different sets of social abilities. In the next section we will rephrase these assumptions into hypotheses.

3.3 Hypotheses

The aim of this user study is first of all to establish how the UTAUT performs for a robot and a screen agent, both used in different conditions. Furthermore, we study the effect of social abilities in two assistive social robot systems on their acceptance by elderly users. In these specific experiments, we want to measure acceptance both as (a) functional acceptance by using the UTAUT acceptance model and (b) social acceptance by using additional questions and observations.

Our first hypothesis concerns the issue of implemented and perceived social abilities. As implemented abilities will not necessarily be perceived, we hypothesize that in these experiments a condition in which more abilities are implemented will evoke a higher score on perceived social abilities:

(1) An assistive social robot in which more advanced social abilities are implemented will be perceived to be more sociable by its users.

The second hypothesis relates to our aim to find a model that predicts acceptance accurately. In technology acceptance modeling this accuracy of prediction is usually expressed in the percentage of the explained variance in the Intention to Use the system that is predicted by the determining constructs (Lee et al. 2003; Venkatesh et al. 2003). A model that results in a score of 50% is considered sufficiently accurate.

(2) The UTAUT model will be able to adequately explain the variance in the Intention to Use the system.

Our third hypothesis focuses on the difference that we expect as a result from more advanced social abilities:

(3) The implementation of more advanced social abilities in an assistive social robot will lead to a higher score on acceptance of the robot by elderly users.

Moreover, we expect that this difference will be independent from the embodiment:

(4) The influence of sociable abilities on acceptance will be similar for different embodiments.

In the demands we established for an accurate model in section 1.5, we stated that it should have the ability to explain acceptance under a wide variety of circumstances. This means it should be able to explain acceptance equally well for more and less sociable conditions in the different systems:

(5) The UTAUT methodology will have an equal explanatory power for all used systems and conditions.

3.4 Used systems

We chose a (physically embodied) robot and a screen agent that were very different systems, although they had comparable functionalities as we will describe in section 3.6. Both systems have already been described in summery in the previous chapter (see Table 2.1 and Appendix A). In this section we will describe them in more detail and discuss the social abilities that were implemented during the creation of the more and less sociable conditions.

3.4.1 iCat

The iCat is basically a service type robot, although it has some companion type capabilities. It is made of hard plastic and has a cat-like appearance, with movable lips, eyes, eyelids and eyebrows to display different facial expressions to simulate social and emotional behavior.

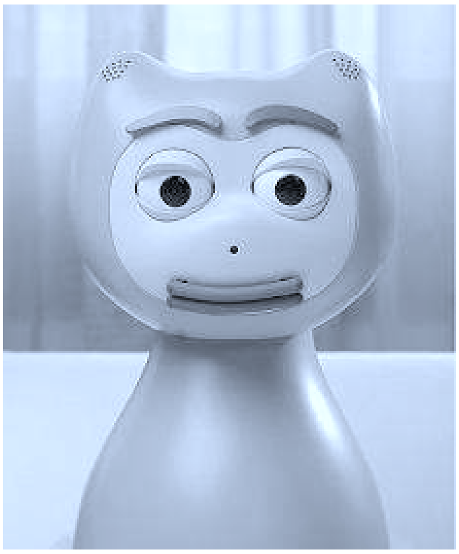


Figure 3.1. The iCat

There is a camera installed in the iCat's nose which can be used for different computer vision capabilities, such as recognizing objects and faces. The iCat's base contains two microphones for sound input and a loudspeaker is built in for sound output. Its design aim is to be a research platform for human-robot interaction, possibly in an intelligent home environment. Athough it is primarily a robot and potentially operates autonomously, studies typically investigate how users respond to the iCat as an (affective) interface to new technology (Bartneck et al. 2004; De Ruyter et al. 2005; Markopoulos et al. 2005; van Breemen et al. 2005; Looije et al. 2006; Pereira et al. 2008).

To generate speech we used a text to speech engine with a Dutch female, speaker. Figure 3.3 shows how we used this robot in a group session and in an individual test session.

3.4.2 Annie

Annie is a service type screen agent with companion type features, developed for our tests by students of the Hogeschool van Amsterdam, University of Applied sciences (HvA), in the Netherlands. The students made use of Chartoon software[1]. It features a female character being able to move the same facial parts as the iCat (lips, eyes/eyelids, eyebrows). It was used on a 17 inch LCD screen in combination with a webcam (attached to the screen), a desk microphone and two speakers.

Figure 3.2. Screen agent Annie

3.4.3 Embodiment of social abilities for the used systems

To create a more and less sociable condition for each system, we have to identify social skills that can be implemented in the used systems. This means we have to link the field of social psychology, where the evaluation of social skills is more common, to the field of HRI.

A tool that is commonly applied to assess social abilities is Gresham & Elliott's Social Skills Rating System (SSRS) (Gresham and Elliot 1990). It is often used in social research involving children and students as participants (Gresham et al. 1998; McClelland and Morrison 2003) The SSRS describes five basic clusters of social behavior on which tested persons can be rated:

1. Cooperation - behaviors such as helping others, sharing materials and complying with rules.
2. Assertion - initiating behaviors such as asking others for information or proactively introducing oneself and behaviors that are responses to others' actions such as responding to peer pressure.
3. Responsibility - behaviors that demonstrate the ability to communicate with concern for the property and wellbeing of others.
4. Empathy - behaviors that show concern for someone's feelings.
5. Self-control - behaviors that emerge in conflict situations such as responding appropriately to teasing or to corrective feedback.

These clusters match the aspects found in Human-Robot Interaction literature on social robots and agents, although the emphasis differs with the robot type and tasks. Describing Kismet, for example, Breazeal describes cooperation and

[1] http://old-www.cwi.nl/projects/FASE/CharToon/

related skills by creating the possibility to let the robot use attentional cues to recognize task-relevant events and objects and empathy related skills with affective expressions (Breazeal 2000; Breazeal 2002). In later work concerning the more sophisticated robot Leonardo, she adds skills related to assertion by enabling the robot to show enthusiasm about the actions it can perform and to be aware of its success and failures (Breazeal 2003).

The concept of responsibility is harder to find in HRI. The term is used as a characteristic of social behavior by De Ruyter et al. (2005), but not clearly defined. Hinds et al. use the term responsibility, but in their study it is rather an aspect of tasks that can be shared between humans and robots. Furthermore Dautenhahn (2004), describes behavior that can be characterized as taking responsibility (without using the term) when describing the characteristics robots can develop by education. If we change our scope to interaction with intelligent systems in general we find a study with high school students by Tzeng (2004) that describes responsible computers that apologize. The study shows that this has a positive influence on user experience, although it does not affect users' performance.

Similar constructs also appear to be relevant abilities in related studies (Fong et al. 2005; Gockley et al. 2006; Looije et al. 2006). In comparable studies, like the experiments with a robot and screen agent by Shinozawa et al. (Shinozawa et al. 2003; Shinozawa et al. 2004) and the experiment with iCat by De Ruyter et al. (De Ruyter et al. 2005) Trust and Competence are also found to be relevant concepts in human-agent interaction. They impact the appreciation of a robot (or screen agent) and the rating of its social abilities.

This leads to the following social abilities to generally be relevant in human-agent interaction:

1. Cooperate. Actually this is a set of abilities that ensures smooth interaction and collaboration, like turn taking in a dialogue, giving and receiving feedback (also using eye contact) and being adaptive to the wishes of the conversational partner (Breazeal 2000; Breazeal 2002; Breazeal 2003).
2. Express empathy. A robot is able to demonstrate that it knows what positive and negative experiences are to a human – even if it cannot truly 'feel' (Breazeal 2000; Breazeal 2002; Breazeal 2003).
3. Show assertion-related behavior. This can simply mean that a robot acts proactively, introducing itself and invite the user to interact. However, this has to be implemented cautiously, since too much proactive behavior by a robot could cause anxiety (Breazeal 2003; Nomura et al. 2006).
4. Exhibit self control. For a robot this could mean to show awareness of its successes and failures (Breazeal 2003; De Ruyter et al. 2005).
5. Show responsibility. For a robot this typically means that it shows that it has no intention to do any harm and that it wants to perform well (which means it apologizes for any mistakes) (Dautenhahn 2004; Tzeng 2004; De Ruyter et al. 2005).

6. Gain trust. A robot's behavior should install trust (Shinozawa et al. 2003; De Ruyter et al. 2005; Shinozawa et al. 2005).
7. Show competence. A robot should demonstrate its capability to do what it is supposed to do (Shinozawa et al. 2003; De Ruyter et al. 2005; Shinozawa et al. 2005).

The issue of identifying social abilities that can be embodied in a robot is also addressed in a study by Lee et al. (2006). It attempts to list social abilities referring to Sternberg et al.'s prototypes of socially competent behavior which in its turn is based on "lay people's observations" (Sternberg et al. 1981). Although this attempt results in a list that is very different from the list above, it faces the same problems in the process to translate the concept of social abilities into specific acts that can be mapped onto a robot's body (in case of Lee et al. the subject is an android). This matter is addressed by Lee et al. as the criterion of 'embodifiability': only limited acts can represent a specific ability and some abilities cannot be represented because of the limitations of the specific agent or the context. We will now face this issue of embodifiability concerning the above mentioned seven abilities.

Lee et al. (2006) present a list of 27 behavioral codes with operational definitions that pass their criterion of embodifiability. Although it is directed towards the specific category of androids, it covers the possibilities of embodiment of the above listed social abilities. With the specific agents (not mobile, limited possibilities of body movement) and the context of our experiments (just a short time to get acquainted and a very limited set of tasks to perform) we found only the following five behavioral features to be applicable (the numbers refer to the above listed abilities):

- listening attentively, for example by looking at the participant and nodding (1, 2),
- being nice and pleasant to interact with, for example by smiling (1, 2, 7),
- remembering little personal details about people, for example by using their names (6, 7),
- being expressive, for example by using facial expressions (2, 3),
- admitting mistakes (2, 5, 6).

Most items of the above mentioned list are more or less represented in the behavioral features - only the feature 'exhibit self control' (4) is not represented at all. We could find no acts that we would be able to embody and that could represent this feature.

Applying these acts for both the iCat and Annie, the difference between the more social and less social condition was realized with the following behavioral features:

- The robot in the more social condition would gaze straight at the conversational partner; the robot in the less social condition would just gaze in the distance, past the participant.

- The robot would make a mistake by saying ‘good morning’ when it would be afternoon or the other way round. When this would be made clear, the robot in the more social condition would admit this mistake and apologize, the robot in the less social condition would not.
- The robot in the more social condition would smile when appropriate (showing empathy, gaining trust), for example when greeting or when a participant would say something funny and express cheerfulness in its movements (moving the eyebrows and eyelids); the robot in the other condition did not do this.
- The robot in the more social condition remembered the participant’s name and would use it when addressing the participant; the robot in the less social condition would not.
- The robot in the more social condition would acknowledge the conversation partner by nodding and blinking; the robot in the less social condition would not do this.
- The robot in the more social condition would be better in turn taking by waiting until the conversation partner finished speaking; the robot in the less social condition was less verbally polite (would be interruptive).

3.5 Pilot study: lessons learned

As we will discuss in section 3.6, we used for our experiments a Wizard of Oz setting in which the robot is controlled by a hidden operator. The participants were to use the robot that was presented to them as autonomous for a few minutes.

Because we had no experience with setting up field experiments with older adults and it was also our first time to use the Wizard of Oz setup, we first organized a pilot session with 28 participants in an eldercare institution in the Dutch city of Almere. The participants were to interact with the iCat robot for a few minutes and fill out a questionnaire afterwards. This would enable us to learn about organizational factors and have a first impression of the responses of elderly users to a robot and their behavior during an experiment.

The experiment made it very clear that we had indeed a lot to learn: due to setbacks and organizational mistakes we obtained usable data of only 11 of the 28 participants.

First, there were organizational issues, due to our inexperience with setting up an experiment cooperating with the nursing staff. For example, we asked staff to invite participants and expected them to be present in the waiting room at the start of the experiments. But as soon as we were ready to let the participants into the testing room, there appeared to be no one waiting. We had to get them out of their apartments ourselves, which took a lot of time, also because some participants were not dressed yet. Also, a lot of participants came during an earlier or later session than the one they were invited to. This was mostly because somewhere between the invitation and the actual experiment, they forgot what time it would start and they asked someone who was also invited, but at another time.

Second, there were issues concerning the mental state of the participants that we took too little into account: about half of the participants had forgotten about the experiment and many of the remaining half had forgotten what it was about. Moreover, some participants forgot during the experiment what it was about, just a few minutes after we had explained it. Furthermore, for many participants the questionnaire was longer than their memory of the session lasted. Many of the participants had trouble remembering their experience when the experiment was 15 minutes or longer ago.

Third, there were behavioral issues that we did not take into account due to inexperience in dealing with groups of elders: some participants refused to work on the given task with the robot; they simply started a conversation with it, ignoring all instructions. Also, some participants walked away as soon as it was time for the questionnaire, because they did not find answering questions a necessary thing to do.

Finally, we found many participants thought we were trying to sell the robot, even after we explained that this was not a sales presentation. Later we learned that the room we used was indeed often used for sales presentations. Some participants left because of this, because the robot was too expensive for them. We could not convince them that it was not our intention to sell anything.

Considering our experiences we recognized the following issues as being crucial to successfully set up an experiment in an eldercare environment to gather user experience data. They served as guiding principles in the setup of all further exzperiments.

1. Collecting user experience data in an eldercare environment
We succeeded in collecting only a small amount of user experience data. We learned that collecting a substantial amount of data demands a very strict organization.

2. Ensuring cooperation and participation
The participation of caregivers is essential. They are the ones who know the different participants and how to ensure their participation. We needed them not only to bring the participants to the experiment, but also to stay with them while they were waiting.

3. Selection of participants
Aging adults who are suffering from some degree of dementia can in many cases participate in an experiment like ours, but if they have forgotten their experiences by the time they are questioned about it, this might lead to unreliable data. If these participants are identified before the experiment, it

remains possible to use other methods to gather data on their experiences. If the questionnaire is essential, like in our case, only participants that will remember their experiences long enough should be selected.

4. Communication with participants
Participants have to be well informed about the purpose and procedures both before and during the experiment. They have to be aware that they are participating in an experiment and that a questionnaire is part of the protocol. Furthermore, participants need to have a written invitation, clearly stating the date and time of their attendance.

5. Limiting questionnaires
There is a limit to the length of a questionnaire elderly participants have patience for, even if their memory does not fail them. Although there are of course differences between individuals, a questionnaire containing up to 30 questions is about as much as most elderly test participants can take.

6. Anticipating non procedural behavior
Many participants may express demands that are not appropriate to a robot's functionalities. This should be anticipated by having standard replies like 'I am sorry, but I am not programmed to do this'.

3.6 Experimental methods

The experiments from which the data were used will be referred to as 'Experiment 1' for the study with the iCat robot and 'Experiment 2' for the study with screen agent Annie.

These experiments were carried out at eldercare institutions in Lelystad and Almere, the Netherlands. Experiment 1 took place in November and December 2005, Experiment 2 in May 2006

3.6.1 Subjects

Experiment 1 and 2 each featured a group of 42 participants. For each of these experiments, participants were elderly inhabitants of the institutions, living more or less independently, or needing daily care and who volunteered for the study. From the data some participants were excluded because of disturbances during the observation session and severe hearing problems. Furthermore, to avoid user bias, we did not want participants that had participated in the experiment with the iCat to take part in the experiment with the screen agent. There were, however, a few participants of Experiment 1 present at Experiment 2. We still allowed them to take part in the experiment, but excluded them from the results of Experiment 2. Table 3.1 shows the age and computer experience of the participants. The age ranged from 65 to 94; the average age was 78.4 for

Experiment 1 and 79.5 for Experiment 2. For computer experience the answer could either be 'yes' (scores 1) or 'no' (scores 0) – of the 36 participants in Experiment 1 there where 13 with computer experience and for the 33 participants in Experiment 2 there were also 13.

System	Item	N	Minimum	Maximum	Mean	Std. Deviation
iCat	Age	36	65	92	78.36	7.82
	Experience	36	0.00	1.00	0.32	0.45
Annie	age	33	65	94	79.48	7.64
	exp	33	0.00	1.00	0.39	0.46

Table 3.1 Age and computer experience of participants in experiments 1 and 2

When asking the nursing staff to recruit participants, we asked them to avoid those whose mental condition was prejudicial to filling in a questionnaire. Otherwise there was no selection on mental or physical health features.

3.6.2 Procedure

A specific interaction context was created where the robot could be used in a Wizard of Oz fashion which we will discuss in section 3.5.5. The participants were first exposed to the robot in groups (8 participants per group). After a short introduction by one of the researchers the robot told them what its possibilities were: an interface to domestic applications, monitoring, companionship, information providing, agenda-keeping and memorizing medication times and dates. They were told that for today's experiment, the robot was only programmed to perform three tasks: setting an alarm, giving directions to the nearest supermarket and giving the weather forecast for tomorrow. The experimenter subsequently demonstrated how to have a conversation with the robot in which it performed these tasks.

Figure 3.3. robot experiment: group session and individual session

After this group session (Figure 3.3 left), the participants were invited one by one to have a conversation with the robot ((Figure 3.3 right), while the other group members were waiting in a different section of the room. The conversation was standardized (by task related scripts) as much as possible and we asked the participants to have the robot perform the three simple tasks. While being engaged in conversation, the participants' behavior was observed by a researcher and recorded by camera. The group session and the individual sessions lasted both about 5 minutes, so the maximum time spent with the robot was 10 minutes for each participant.

Immediately after the individual session, the participants would be taken to a room where they were asked to fill out the questionnaire.

3.6.3 Questionnaire

Before the pilot experiment was carried out, the questions concerning acceptance were adapted from the UTAUT questionnaire (included in Appendix D) and tested on a group of ten elderly. Of this test group, three participants had difficulty indicating the level to which they agreed with statements. When we rephrased the statements to questions, they responded far better, so we decided to use questions instead of statements. Also, because some of the participants had trouble reading, it turned out to be much easier for most of them if they were asked the questions by an interviewer. The answers to the questions were given on a five point scale with 1 being the negative end of the scale and 5 being the positive end of the scale.

Furthermore, since UTAUT was developed for using technology at work, the questions needed to be adapted to a domestic user environment. Questions that could not be adapted were omitted.

The final questionnaire contained 27 items of which 19 were part of the UTAUT-derived constructs; each construct was represented by at least two questions:

- Performance Expectancy (PE) was represented by two questions which were adapted to a non working environment. Two original UTAUT questions were removed because they were too much related to a working situation (increase productivity and getting a raise) to be adapted.
- Effort Expectancy (EE) was represented by three questions. They were adapted to using a voice operated system.
- Social Influence (SI) was represented by two questions.
- Attitude toward using technology (AT) was represented by two questions. Two original UTAUT questions were removed because they were too much related to a working situation.
- Self-efficacy (SE) was represented by three questions. One original UTAUT question was removed because it was too much related to a working situation.
- Anxiety (ANX) was represented by two questions. Again, two original UTAUT questions were removed because they were too much related to a working situation.
- Intention to Use (ITU) was represented by the original three UTAUT items that were translated into Dutch questions as literal as possible.

We also created a manipulation check by adding five questions (also to be answered on a five point scale) concerning Perceived Social Abilities (SA) so that we could establish whether the Perceived Sociability matched the different conditions.

Code	Question
CE	1) Have you ever used a computer?
CE	2) Do you still sometimes use a computer?
CA	3) Did you feel uncomfortable talking to a robot?
PE	4) Do you think the robot would be useful to you?
PE	5) Do you think the robot would help you do things?
EE	6) As you have noticed, you control the robot by speech. Do you think you can easily communicate with it that way?
EE	7) Do you think you can quickly learn how to control the robot?
EE	8) Do you think the robot is easy to use?
SI	9) Do you think many people would be pleased if you would have the robot?
SI	10) Are these people whose opinion you value?
SI	11) Are these people who are important to you?
SI	12) Do you think the staff would be pleased if you would have the robot?
SA	13) Did you find the robot a pleasant conversational partner?
SA	14) Would you consider the robot to be social?
SA	15) Would you trust the robot if it gave you advice?
SA	16) Would you follow the robot's advice?
SA	17) Do you feel understood by the robot?
AT	18) Do you think it is a good idea to use the robot?
AT	19) Would you like to use the robot?
SE	20) Do you think you could work with the robot without any help?
SE	21) Do you think you could work with the robot if you could call someone for help?
SE	22) Do you think you could work with the robot if you had a good manual?
ANX	23) Do you feel at ease with the robot?
ANX	24) If you were to use the robot, would you be afraid to make mistakes or break something?
ITU	25) If you could have the robot, would you want it immediately?
ITU	26) If you could have the robot, would you want it in a few months?
ITU	27) If you could have the robot, would you want it in a few years?

Construct codes:

UTAUT:	PE	performance expectancy
	EE	Effort Expectancy
	SI	social influence
	AT	attitude toward using technology
	SE	self-efficacy
	ANX	anxiety
	ITU	Intention to Use
Other:	CE	computer experience
	CA	conversational acceptance
	SA	social abilities

Questions are translated – the list used in the experiments is in Dutch.

Table 3.2. The questionnaire on acceptance as used in the experiments

Furthermore, we added two questions on experience with computers (CE) (to be answered with yes or no) and one question concerning the extent to which people felt (un)comfortable when talking to a robot (CA) (to be answered with 'yes', 'no' or 'a bit'). This means that there were three constructs added for this special context: 'social abilities', 'conversational acceptance' and 'computer experience'. The issue of conversational acceptance was only represented by one question in the questionnaire (code CA), but was also measured extensively by our observation model (see next section).

Table 3.2. gives a list of the questions, as well as a categorization whether a construct is a UTAUT-derived construct or not.

3.6.4 Behavior observation

We found it important to look for an instrument that would make it possible to detect differences in user response that could be linked to the differences in perceived social abilities and that could be used in addition to the (adapted) UTAUT questionnaire. Besides the control questions (CA and SA in Table 3.1.) we wanted to apply user behavior observation and we videotaped the individual sessions.

The instrument we developed for this was a list of items considering conversational expressiveness: non verbal behavior that conversational partners show and that is known to be more intense when users are enjoying a conversation and feel more involved in it (Cappella 1983; Coker and Burgoon 1987). The list was generated by listing classical feedback gestures that are often observed in interaction research. An example of this is a study of gestures related to verbal feedback expressions by Allwood and Cerrato (Cerrato 2002; Allwood and Cerrato 2003) which involves collection of these gestures that are also applied to human machine communication. Also a study by Heylen et al. (Heylen et al. 2006) on detecting what people feel when they communicate with machines, addresses feedback gestures that are listed below. Although these gestures are often linked to specific communicative functions, we decided to leave them free from interpretation and simply count their appearances. The gestures are not specifically intentional or non intentional, but they can be identified as conversational behavior.

1. *Nodding head*
2. *Shaking head*
3. *Greeting*
4. *Lifting shoulders*
5. *Suddenly moving away*
6. *Suddenly approaching*
7. *Smile*
8. *Laugh*
9. *Raise eyebrows*
10. *Frown*

From this list, the numbers 1, 3, and 5 to 10 are also used by Cerrato (2003) and 1, 2, 7, 9 and 10 by Heylen et al. (2006). The only gesture that we did not find in studies of human machine interaction studies is number 4. *Lifting shoulders*. We included this gesture simply because our participants executed it a few times. Furthermore, we added the behavior of verbal greeting to it, because we considered this also a sign of relational feedback.

The sessions were recorded by video and were analyzed afterwards. During analysis the items of conversational expressiveness were counted for each participant by two different observants who were unaware of the different conditions of the robot.

3.6.5 Wizard of Oz

Both systems were used in a Wizard of Oz setup, which means that a hidden operator controlled the robot. This setup is often used for prototype testing and (especially) user response research (Green and Wei-Haas 1985; Francony et al. 1992; Dahlbäck et al. 1993), either to test functionalities that are not (yet) fully realized in the tested system, or to have more control over the experimental conditions. For our experiments, both were applicable: the Wizard of Oz set up made it possible to avoid problems with voice recognition and enabled us to have the individual sessions follow roughly the same interaction patterns.

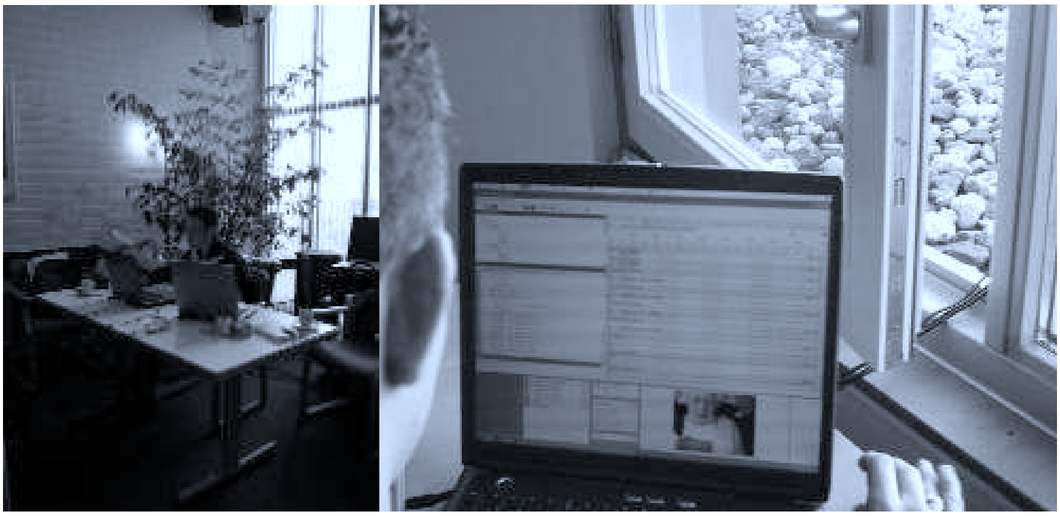


Figure 3.4. Hidden operator during the iCat experiment

Figure 3.4 shows the hidden operator and his screen to operate the iCat and Figure 3.5 shows these for the screen agent. In both settings, the operator had a set of pre-programmed acts (combinations of speech and head movements) available, form which he could choose during the conversation, either to proactively take control of the conversation ('Please state what you want me to do') or to reply to the participant's remarks. He was able to view the participant through a small camera, which was hidden in the nose of the iCat and on top of the screen for Annie. He could hear the participants through a small flat microphone that was attached to the participant's table.

Figure 3.5. Setup and hidden operator during the screen agent experiment

3.7 Results of experiments 1 and 2

From the 42 participants for Experiment 1, we obtained usable results from 36 users. The participants that were omitted from the results had not completed their questionnaire. From the 42 participants for the screen agent experiments, we obtained questionnaire results from 33 users. We omitted 9 participants from the results, either because they had not completed their questionnaire or because they had already participated in the experiments with the robot. Three participants that had not completed their questionnaire were included in the results for behavior observation.

The descriptive statistics in Table 3.3a and 3.3b show that the scores are close for several constructs, both for systems and conditions. The high scores on Effort Expectancy and Anxiety actually denote the systems and conditions were considered easy to use and there was little outspoken anxiety - the scores for the negative Anxiety item (24) were reversed before they were processed (a high value for the variable means low anxiety).

	All		Experiment 1		Experiment 2	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Computer Experience	0.355	.455	0.319	.450	0.394	.464
Conversational Acceptance	2.725	.539	2.722	.513	2.727	.574
Performance Expectancy	3.783	1.126	3.694	1.130	3.879	1.132
Effort Expectancy	4.116	1.049	4.083	1.150	4.152	.943
Social Influence	3.652	.975	3.681	.880	3.621	1.083
Social Abilities	3.623	.899	3.633	.881	3.612	.933
Attitude	3.667	1.444	3.750	1.500	3.576	1.398
Self-efficacy	4.188	1.014	4.361	1.099	4.000	.890
Anxiety	4.362	.907	4.236	1.004	4.500	.781
Intention to Use	3.285	1.49	3.398	1.547	3.162	1.441

Table 3.3a Descriptive statistics for experiments 1 and 2

The first step in our analysis is to compare the effect of the two conditions on the acceptance of both agents. Then we will study correlations between constructs. Our focus in this is to find out whether a higher perception of social abilities correlates with a higher score on other constructs.

	More social		Less social	
	Mean	St. Dev.	Mean	St. Dev.
Computer Experience	0.288	.451	0.417	.455
Conversational Acceptance	2.909	.384	2.556	.607
Performance Expectancy	4.030	.910	3.556	1.264
Effort Expectancy	4.253	.936	3.991	1.142
Social Influence	3.773	.902	3.542	1.038
Social Abilities	3.849	.689	3.417	1.023
Attitude	3.833	1.407	3.514	1.481
Self-efficacy	4.278	.846	4.107	1.152
Anxiety	4.424	.821	4.306	.988
Intention to Use	3.505	1.444	3.083	1.525

Table 3.3b Descriptive statistics for the conditions

Table 3.3b shows that Computer Experience was higher for participants that had interacted with the less social condition. This should be taken into account if there are differences between participants that can be attributed to or are influenced by Computer Experience.

3.7.1 Questionnaire

Before analyzing the scores for the more and less socially communicative conditions, we calculated Cronbach's Alpha for the UTAUT-derived constructs to check internal consistency. Generally in Social Science, an alpha of 0.7 and higher is considered acceptable, (Decoster and Claypool 2004).

Construct	iCat	Annie
Performance Expectancy	.765	.881
Effort Expectancy	.861	.726
Social Influence	.300†	.567†
Social Abilities	.747	.786
Attitude	.889	.780
Self-efficacy	.894	.725
Anxiety	.430†	.159†
Intention to Use	.895	.839
all UTAUT questions	.935	.927
all questions	.924	.933

†Internal consistency for construct below threshold.

Table 3.4. Cronbach's alpha on all constructs for both systems

The construct scores were formed by joining and averaging the scores for the questions that represented it. An exception to this was the SI construct: it was represented by four questions, but two (questions 10 and 11) are dependent questions so only the first and last one were incorporated in the scores. As Table 3.4 shows, the scores on the constructs for Social Influence and Anxiety were too low for both systems, implying that we should not take these constructs into account because of low internal consistency.

For both systems we analyzed the differences between the conditions on the scores for the constructs, using the Mann Whitney U-test, which is common for comparing groups of this magnitude (20 or less per group)(Cohen 1992). The results showed that in fact none of the UTAUT-derived constructs showed a significant difference for either of the systems.

Also the scores on the five questions related to social abilities (SA) did not show any significant differences for the two conditions for either system and neither for the combined scores of both systems. Nevertheless, there is a pattern of higher score for the more social condition as is shown in Figure 3.5 and 3.6.

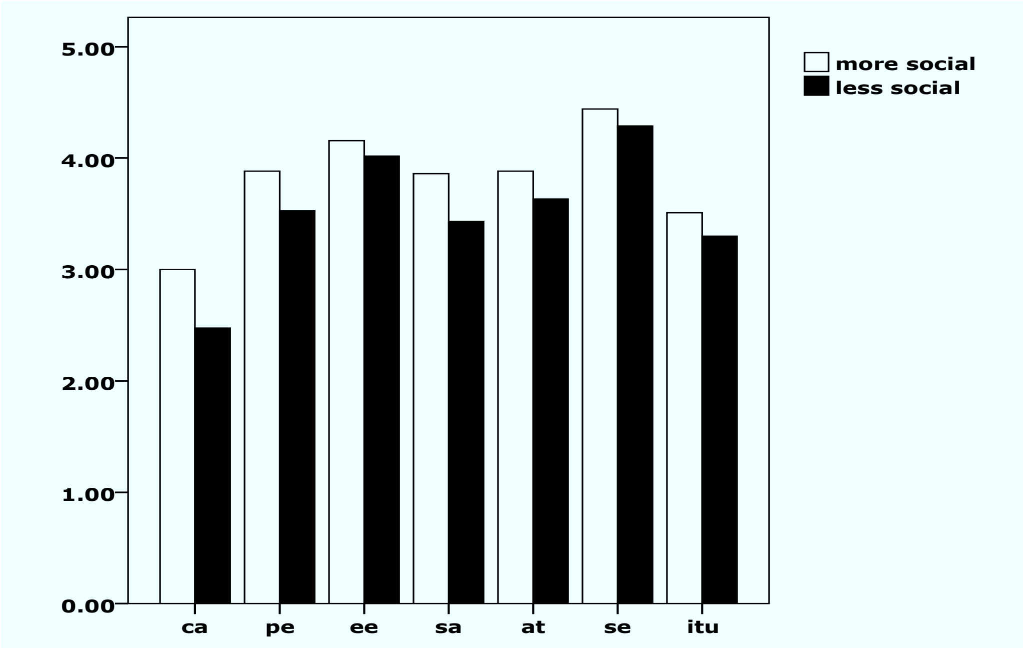


Figure 3.5. Mean scores for the constructs for the iCat

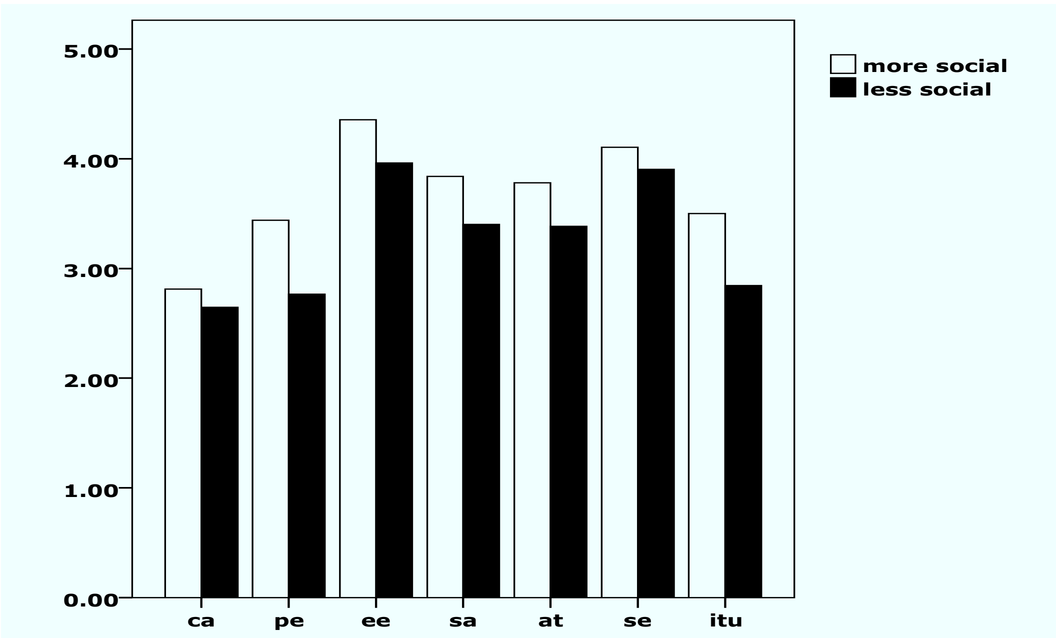


Figure 3.6. Mean scores for the constructs for Annie

A significant difference was found between the two conditions of the iCat on the question 'Did you feel uncomfortable talking to a robot' (question 3 in Table 3.2., related to 'conversational acceptance') which could be answered with 'yes', 'a little' or 'no' (so this is a question with answers on a 3-point scale). All (17) participants who experienced the more social condition reported to feel comfortable (or 'not uncomfortable') about it, while 47% of the (19) participants that encountered the less social condition reported to feel a little or very uncomfortable. Table 3.5 shows how a Mann-Whitney U-test indeed resulted in a significant difference. Table 3.6 shows that this difference did not occur in the results for Annie.

Condition	N	Mean	Mann-Whitney U	Sig. (2-tailed)
more social	17	3.00		
less social	19	2.47	85.0	0.001**

Table 3.5. Score conversational acceptance Experiment 1 (iCat)

Condition	N	Mean	Mann-Whitney U	Sig. (2-tailed)
more social	16	2.81		
less social	17	2.65	114.5	0.277

Table 3.6. Score conversational acceptance Experiment 2 (Annie)

Comparing the results of the iCat to those of Annie, we found no significant differences between the scores for the constructs (neither for the more social, nor for the less social condition). For the individual questions we also did not find any significant differences except for question 24. On the question if they would be afraid to make mistakes or break something the score for the iCat was much higher for both conditions (see Table 3.7 – note that this concerns a reversed score: more negative response, i.e. being less afraid, scores higher).

Experiment	N	Mean	t	Sig. (2-tailed)
1 (iCat)	36	4.06		
2 (Annie)	33	4.82	-3.031	0.004**

Table 3.7. T score question on being afraid to make mistakes or break something comparing both systems

Our conclusion is, that despite this pattern of higher scores for the more social condition, there is no significant difference for any of the UTAUT constructs.

Source	Dependent Variable	F	Sig.
robot	Conversational Acceptance	.003	.955
	Performance Expectancy	.461	.499
	Effort Expectancy	.075	.785
	Social Abilities	.015	.903
	Attitude	.246	.622
	Self-efficacy	2.185	.144
	Intention to Use	.413	.523
condition	Conversational Acceptance	7.834	.007*
	Performance Expectancy	3.125	.082
	Effort Expectancy	1.081	.302
	Social Abilities	4.033	.049*
	Attitude	.845	.361
	Self-efficacy	.521	.473
	Intention to Use	1.440	.235

Table 3.8. Two-way ANOVA for all reliable constructs

To further explore the differences between conditions and systems, we performed a two-way ANOVA on the joint scores for both systems (no interaction effects between *robot* and *condition* were found). In Table 3.8 the results are presented. This analyses shows a significant score for Conversational Acceptance and Social Abilities where the conditions are compared. The latter is remarkable, since the difference between the conditions for this construct was not significant in the

Mann-Whitney test for either of the systems. Nevertheless, we can take this as an indication that overall the more social condition would be indeed recognized as such, which confirms our first hypothesis.

3.7.2 Behavior observation

As described in section 3.5.4, the recorded video's of the individual sessions were all analyzed by two different observers (who were unaware of the different conditions of the robot) to measure conversational expressiveness. To examine inter-rater agreement we calculated Cohen's kappa(Cohen 1960), which is a common procedure if there are just two raters (Banerjee et al. 1999). The results in Table 3.9. show an average score of 7.65, which means there is substantial agreement (Landis and Koch 1977).

Behavior	κ	Behavior	κ
nodding	0.665	laugh	0.820
shaking head	0.788	raise eyebrows	0.679
greet with hand	0.866	greet with words	0.806
lifting shoulders	0.605	frown	0.727
move away	0.876	Total counts	0.732
approach robot	0.681		
smile	0.929	Average	0.765

Table 3.9. Cohen's kappa for behavior observation items

Figures 3.7 and 3.8 show that there are remarkable differences between the scores for the two conditions and there is a certain pattern of more expressiveness in the more social condition.

Agent:	iCat t	Sig.	Annie t	Sig.	Combined t	Sig.
positive	2.450	0.020*	2.017	0.052	2.902	0.005**
negative	-0.986	0.333	0.457	0.650	-0.471	0.639
all items	2.063	0.047	2.024	0.051	2.607	0.011*

Table 3.10. Conditions of both systems: t scores on categorized behavioral observations

If we look at the total number of times a type of behavior (positive/negative) occurred for the different conditions (Table 3.10.), there is a significant difference both in total expressions and in the total amount of expressions that can be categorized as positive expressions (all behaviors except shaking head, move away and frown).

As with the questionnaire results we performed a two-way ANOVA on the combined scores. Table 3.11 shows the results. The analysis shows that there are no significant differences between the two systems and a few differences between the conditions concerning nodding and smiling. The condition scores show the same pattern as the t-tests in Table 3.10: a significant difference both in total expressions and in the total amount of expressions that can be categorized as positive (we found no interaction effects between robot and condition).

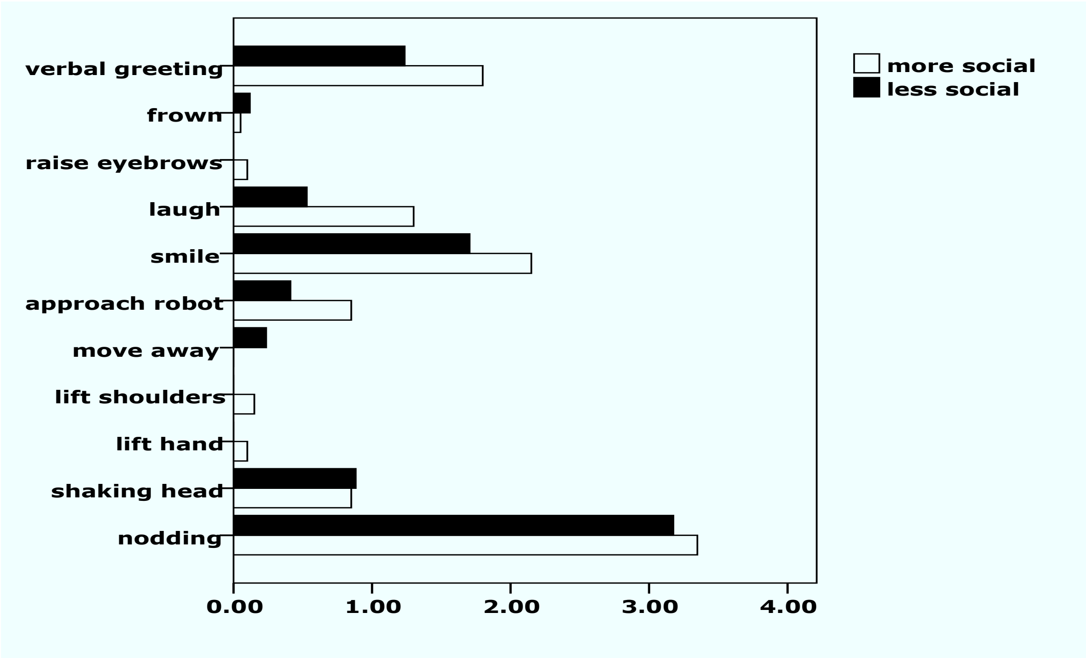


Fig. 3.7. Average counts per participant on conversational expressions for iCat

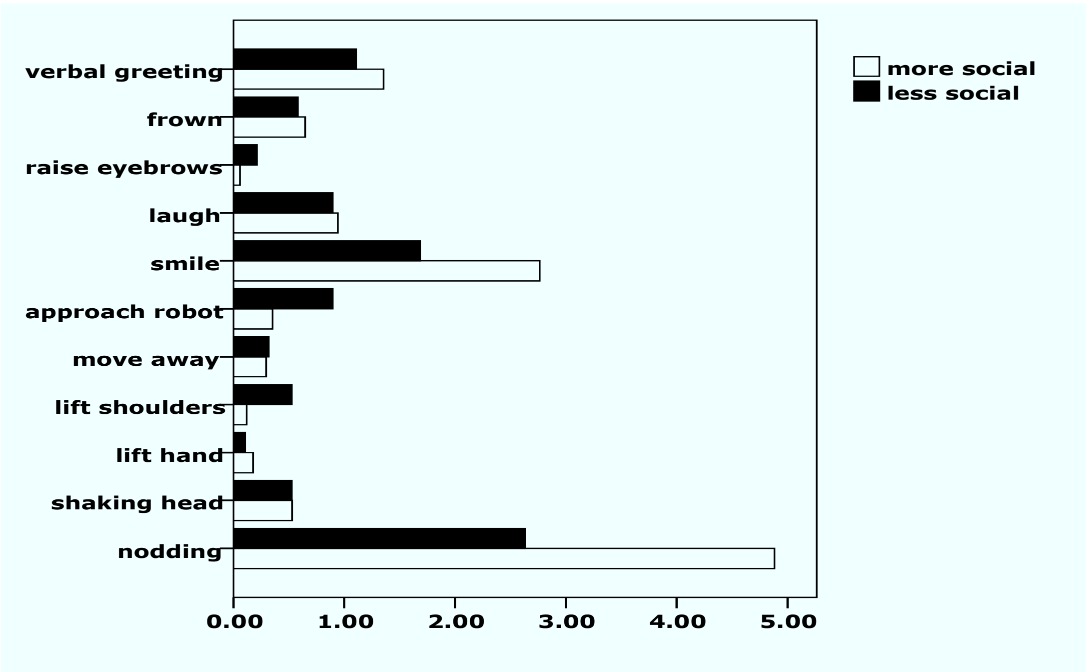


Fig. 3.8 Average counts per participant on conversational expressions for Annie

3.7.3 Correlations and regression

To establish how the scores on constructs interrelated, we performed a correlation analysis, using the scores on UTAUT-derived constructs, the construct of Social Abilities (SA) and Conversational Acceptance. As Table 3.12 shows, the score on the (non-UTAUT) construct of Social Abilities did show correlations with some of the UTAUT constructs and with the question on feeling uncomfortable talking to the iCat.

Source	Dependent Variable	Type III Sum of Squares	Mean Square	F	Sig.
Robot	nodding	4.427	4.427	.732	.395
	shaking head	2.078	2.078	3.313	.073
	greet with hand	.150	.150	1.707	.196
	lifting shoulders	1.108	1.108	2.631	.109
	move away	.637	.637	3.463	.067
	approach robot	.001	.001	.001	.971
	smile	1.597	1.597	.692	.408
	laugh	.000	.000	.000	.990
	raise eyebrows	.130	.130	1.523	.221
	greet with words	1.512	1.512	1.492	.226
	frown	5.085	5.085	16.528	.000**
	positive	15.384	15.384	1.032	.313
	negative	3.468	3.468	3.165	.080
	all	34.925	34.925	2.046	.157
Condition	nodding	26.681	26.681	4.413	.039*
	shaking head	.004	.004	.006	.937
	greet with hand	.133	.133	1.515	.223
	lifting shoulders	.304	.304	.721	.399
	move away	.300	.300	1.629	.206
	approach robot	.049	.049	.072	.789
	smile	10.552	10.552	4.572	.036*
	laugh	3.030	3.030	2.399	.126
	raise eyebrows	.012	.012	.142	.708
	greet with words	2.996	2.996	2.956	.090
	frown	9.79	9.79	.000	.999
	positive	129.239	129.239	8.666	.004**
	negative	.128	.128	.117	.733
	all	123.999	123.999	7.263	.009*

Table 3.11. Two-way ANOVA on the combined scores for behavioral observations

The scores for the screen agent show a significant correlation of Social Abilities with every item (Table 3.13). These results indicate that the question on Conversational Acceptance and the construct of Social Abilities could very well be part of a more accurate model.

Construct	Pearson Correlation	Sig. (2-tailed)
Conversational Acceptance	-0.337	0.045*
Performance Expectancy	0.210	0.219
Effort Expectancy	0.580	0.000**
Attitude	0.473	0.004**
Self-efficacy	0.264	0.120
Intention to Use	0.201	0.241

Table 3.12. Correlation of perceived social abilities with constructs for the iCat

The UTAUT model states that three constructs directly determine Intention to Use: Performance Expectancy, Effort Expectancy and Social Influence. The influences of other constructs are not direct: they influence Performance Expectancy, Effort Expectancy. Since Social Influence has turned out not to be a reliable construct for either agent (see Table 3.4), we performed a regression analysis with the first two constructs. The results are shown in Table 3.14 in which also the adjusted R^2 value is given – this indicates the percentage of variation of the dependent variable that is explained by the independent variables.

Construct	Pearson Correlation	Sig. (2-tailed)
Conversational Acceptance	0.517	0.001**
Performance Expectancy	0.584	0.000**
Effort Expectancy	0.625	0.000**
Attitude	0.744	0.000**
Self-efficacy	0.522	0.001**
Intention to Use	0.729	0.000**

Table 3.13. Correlation of perceived social abilities with constructs for Annie

In all three of the analyses (iCat, Annie and combined scores) Performance Expectancy turns out to be a determining influence on Intention to Use and Effort Expectancy is never a determining influence. Furthermore, there is a difference in variance explained of .22 between iCat and Annie. This indicates that the model does not predict equally well for both systems, which makes us reject Hypothesis 5. UTAUT turns out to predict acceptance accurately for Annie, but not for iCat (which also means we have to reject Hypothesis 2).

Agent type	Independent	Dependent	Beta	t	Sig.	R^2
iCat	Performance Expectancy	Intention to	.551	2.108	.043*	.37
	Effort Expectancy	Use	.064	.245	.808	
Annie	Performance Expectancy	Intention to	.579	3.515	.001**	.59
	Effort Expectancy	Use	.242	1.472	.151	
Combined	Performance Expectancy	Intention to	.544	3.691	.000**	.45
	Effort Expectancy	Use	.149	1.011	.316	

Table 3.14. Regression analysis for the iCat, Annie and the combined results

	Agent	Independent	Dependent	Beta	t	Sig.	R^2
less social	iCat	Performance Expectancy	Intention to	.701	1.479	.159	.45
		Effort Expectancy	Use	-.034	-.071	.944	
	Annie	Performance Expectancy	Intention to	.832	2.621	.020*	.65
		Effort Expectancy	Use	-.028	-.089	.930	
	Combined	Performance Expectancy	Intention to	.743	2.846	.008*	.52
		Effort Expectancy	Use	-.021	-.081	.936	
more social	iCat	Performance Expectancy	Intention to	.457	1.372	.192	.28
		Effort Expectancy	Use	.099	.298	.770	
	Annie	Performance Expectancy	Intention to	.486	2.349	.035*	.50
		Effort Expectancy	Use	.382	1.850	.087	
	Combined	Performance Expectancy	Intention to	.419	2.290	.029*	.34
		Effort Expectancy	Use	.226	1.232	.227	

Table 3.15. Regression analysis for the conditions

If we look at the regression scores for the separate conditions (Table 3.15), we see that for each system, the variance explained by the coefficient of determination (R^2) is higher for the less social condition. Of course, also the combined scores show this pattern: not only embodiment, but also a more or less social condition has a clear impact on the scores. This does not only indicate that the model does not predict equally well for different conditions, it also indicates that it needs improvement if applied to a more socially interactive system. Moreover, between all these subgroups, the R^2 value varies between .28 and .65, which is a difference of .37.

3.7.4 Evaluation of additional constructs

Agent type	Independent	Dependent	Beta	t	Sig.	R^2
iCat	Performance Expectancy	Intention	.568	2.151	.039*	.38
	Effort Expectancy	to Use	.048	.181	.858	
	Conversational acc.		-.108	-.772	.446	
Annie	Performance Expectancy	Intention	.638	4.133	.000**	.66
	Effort Expectancy	to Use	.001	.005	.996	
	Conversational acc.		.333	2.440	.021*	
Combined	Performance Expectancy	Intention	.549	3.717	.000**	.45
	Effort Expectancy	to Use	.123	.820	.415	
	Conversational acc.		.085	.897	.373	
iCat	Performance Expectancy	Intention	.590	2.209	.034*	.38
	Effort Expectancy	to Use	.122	.448	.657	
	Social Abilities		-.147	-.821	.418	
Annie	Performance Expectancy	Intention	.525	3.310	.003**	.64
	Effort Expectancy	to Use	.116	.688	.497	
	Social Abilities		.283	2.050	.049*	
Combined	Performance Expectancy	Intention	.527	3.513	.001**	.45
	Effort Expectancy	to Use	.115	.734	.465	
	Social Abilities		.080	.694	.490	
iCat	Performance Expectancy	Intention	.450	1.728	.094	.42
	Effort Expectancy	to Use	-.150	-.530	.600	
	Attitude		.383	1.731	.093	
Annie	Performance Expectancy	Intention	.241	1.653	.109	.77
	Effort Expectancy	to Use	.216	1.709	.098	
	Attitude		.551	4.681	.000**	
Combined	Performance Expectancy	Intention	.340	2.344	.022*	.55
	Effort Expectancy	to Use	.015	.105	.917	
	Attitude		.451	3.784	.000**	

Table 3.16. Regression analysis for alternative models, adding Social Abilities, Conversational Acceptance and Attitude

Table 3.16 contains the results of a regression analysis in which an alternative model is explored. First, the question on Conversational Acceptance is added to the two constructs that determine Intention to Use. It shows a higher R^2 value for the individual experiments, but not for the combined scores. Next, the non UTAUT constructs of Social Abilities is added instead and we see the same pattern. Finally, the construct of Attitude is added - this alternative influence is already suggested by some recent studies as explained in section 2.4.3. This shows a higher R^2 for both robots and for the overall scores. The effect of adding constructs is the strongest for Annie, especially for Attitude (.59 to .77).

Table 3.17 contains the results of a regression on the conditions if Social Abilities, Conversational Acceptance and Attitude are added. It shows an improvement of five percent on variance explained in the less social condition and an improvement of 26 percent for the more social condition compared to the original model regression (see R^2 values in Table 3.14). Most of this will be the effect of the addition of Attitude, as Table 3.16 (in the last three rows Attitude is added) shows this to be of strong influence.

Condition	*Independent*	*Dependent*	*Beta*	*t*	*Sig.*	*R^2*
Less social	*Performance Expectancy*	*Intention to Use*	*.674*	*2.484*	*.019**	*.57*
	Effort Expectancy		*-.120*	*-.440*	*.663*	
	Social Abilities		*-.171*	*-.909*	*.371*	
	Conversational acc.		*.017*	*.134*	*.894*	
	Attitude		*.363*	*1.822*	*.078*	
More social	*Performance Expectancy*	*Intention to Use*	*.278*	*1.496*	*.146*	*.60*
	Effort Expectancy		*-.065*	*-.365*	*.718*	
	Social Abilities		*.006*	*.038*	*.970*	
	Conversational acc.		*.260*	*1.767*	*.089*	
	Attitude		*.544*	*3.195*	*.004***	
Both	*Performance Expectancy*	*Intention to Use*	*.352*	*2.414*	*.019**	*.56*
	Effort Expectancy		*.012*	*.084*	*.934*	
	Social Abilities		*-.091*	*-.767*	*.446*	
	Conversational acc.		*.109*	*1.163*	*.249*	
	Attitude		*.480*	*3.812*	*.000***	

Table 3.17. Regression analysis including Social Abilities, Conversational Acceptance and Attitude for the conditions

Agent	*Independent*	*Dependent*	*Beta*	*t*	*Sig.*	*R^2*
iCat	*Performance Expectancy*	*Intention to Use*	*.483*	*1.825*	*.078*	*.45*
	Effort Expectancy		*-.154*	*-.512*	*.612*	
	Social Abilities		*-.114*	*-.581*	*.566*	
	Conversational acc.		*-.118*	*-.759*	*.454*	
	Attitude		*.438*	*1.931*	*.063*	
Annie	*Performance Expectancy*	*Intention to Use*	*.226*	*1.796*	*.084*	*.83*
	Effort Expectancy		*.093*	*.787*	*.438*	
	Social Abilities		*-.038*	*-.316*	*.755*	
	Conversational acc.		*.313*	*3.014*	*.006*	
	Attitude		*.556*	*4.005*	*.000*	

Table 3.18. Regression analysis including Social Abilities, Conversational Acceptance and Attitude for the embodiments

An important effect of the addition of different constructs as shown in Table 3.17 is that the difference between conditions is just three percent. This is despite the difference in dominating constructs: for the less social condition Performance Expectancy is the most dominating influence, while for the more social condition it is Attitude.

Table 3.18 shows the same pattern of increase for the two agents compared to Table 3.14 (eight for iCat and 24 for Annie), but there is still a difference between the two embodiments that is – again, as we have concluded from the results in Table 3.16 – this is mainly due to the addition of Attitude.

3.8 Discussion

It is a remarkable development that our senior citizens might become the pioneers of a new era in which the company of robots becomes as common as the use of cars. They could be among the first to get emotionally close to robotic systems. Our results on conversational acceptance suggest that this closeness will be slightly more probable for three dimensional robots than for screen agents – it might be the case that embodiment is of influence.

We found that while developing autonomous interactive systems like assistive social robots for elderly users it is crucial to work on the implementation of social abilities in order to optimize interaction and acceptance. Our study did show a significant influence of these social abilities, in conversational acceptance and correlations, but remarkably not in scores of UTAUT-derived constructs as was done in other studies that used (an adapted version of) UTAUT as an acceptance model (De Ruyter et al. 2005; Looije et al. 2006). This can be attributed to the relative shortness of the time that our participants were interacting and the simplicity of the tasks. This suggests that it may be necessary to collect data on long-term interaction in which the experience of working with an agent goes beyond the first impression.

Also the simplicity of the tasks might be something to take a closer look at. It enabled us to set up a very strict scenario, allowing few surprises, but it also gave participants less of an impression of what an agent could do for them.

Apart from such differences concerning the experiment itself, there was a difference in the used questionnaire. The other experimenters modified the UTAUT items less then we did and used statements instead of questions, which could have influenced the outcome.

Nevertheless our experiments did show the relevance of looking beyond functional technology acceptance when dealing with autonomous interactive systems and incorporate social acceptance. In addition, it demonstrated how a behavior analysis can complement a questionnaire based model. This is especially the case when dealing with elderly participants, because many of them are difficult to interview, either because of (a) difficulty remembering what happened a moment ago or (b) difficulty focusing on answering a questionnaire longer than a few minutes.

The behavior analysis also lead to what might be one of the most surprising results of our experiment: participants were more expressive when interacting with a more expressive robot or screen agent (though the affect for the screen agent appeared less strong). This type of chameleon effect (Chartrand and Bargh 1999) has also been registered in other research where it could be related to social acceptance (Kahn et al. 2004).

3.9 Conclusions

Regarding our first hypothesis we could confirm that systems in which social abilities were incorporated were recognized as being more sociable (a higher score for the SA construct in the more social condition - see Figure 3.5 and 3.6.).

Our second hypothesis states that the UTAUT model will be able to explain at least 50% of the variation in the Intention to Use the system. We could confirm this for the screen agent, but found this not to be the case for the iCat. However, adding three constructs to the determining influences on Intention to Use showed that it is possible to have a higher predictive strength in this setting.

Our third hypothesis, states that the different systems (robot and screen agent) would both show the effect of more social abilities. This was not the case for the UTAUT constructs. However, there are also some differences between the systems that have to be accounted for. The effect of participants being more expressive when interacting with a more expressive agent appears to be stronger for the iCat than for Annie. Moreover, when dealing with the iCat, participants were more afraid to do something wrong or break something, which might in some cases increase anxiety. This is understandable since the iCat is a physical agent that could indeed be damaged, which is much less the case for Annie.

The fourth hypothesis, stating that more advanced social abilities in an assistive agent will lead to a higher score on acceptance could not be confirmed, since the scores on the UTAUT constructs did not differ significantly. However, the results show a significant correlation of social abilities with determinants of acceptance, indicating that social abilities have some effect.

Our fifth hypothesis states that the model has an equal explanatory power for both systems and both conditions. We have to reject this for both parts of the hypothesis, since we found remarkable differences between the systems and the conditions. The latter suggests that the model is less accurate when applied to a more socially interactive system.

Our experiments suggest that a model that does not only focus on technology acceptance but also on Attitude and Conversational Acceptance and possibly other constructs related to social abilities can be more appropriate when dealing with assistive social robots. A well developed model incorporating these influences can be expected to have a higher score on variance explained than the 37 to 59 percent we found with the original UTAUT model (Table 3.14).

In the next chapter we will focus on the development of such a model.

4. Developing a new model I

Parts of this chapter have been published earlier in (Heerink et al. 2008a; Heerink et al. 2008b; Heerink et al. 2008c; Heerink et al. 2009a)

4.1 Introduction

In the previous chapter we saw how the UTAUT model was not adequate in several aspects. It resulted in constructs that were not reliable; the m odel did not perform equally well for the different systems and conditions, and it could not be used to explain differences in responses to different conditions of the used robot. These differences were related to Conversational Acceptance (although just represented by one question) and Social Abilities, which were not part of the UTAUT methodology. Furthermore, also behavior analysis showed differences. Therefore, we will proceed by developing a more complete and more consistent model incorporating influences that are specific for assistive social robots used by older adults.

In the next section the strategy is described. Subsequently, it is described how the used constructs were established and a preliminary version of the model is presented. After this, research targets are set and experiments to test and validate the model are described and planned.

4.2 Strategy

Our strategy (visualized in Figure 4.1) is as follows:

1. A further evaluation of the results from the experiments with the UTAUT derived model as described in the previous chapter. Reflecting upon these results, we want to establish which constructs are to be used in our new model and how these constructs interrelate (which are determining influences on Intention to Use). Moreover, we will use the results of a principal component analysis to establish whether an alternative, more appropriate grouping of the used questions suggests new constructs.
2. Evaluate constructs presented in recent related research and investigate the possibility of these constructs being a part of our model. For each possible construct we will establish its relevance within the context of assistive social robots and screen agents used by older adults. We will take into account possible overlaps with constructs that have already been chosen as part of our new model. We will hypothesize interdependencies

between the constructs based on indications found by our own analysis and in other studies.

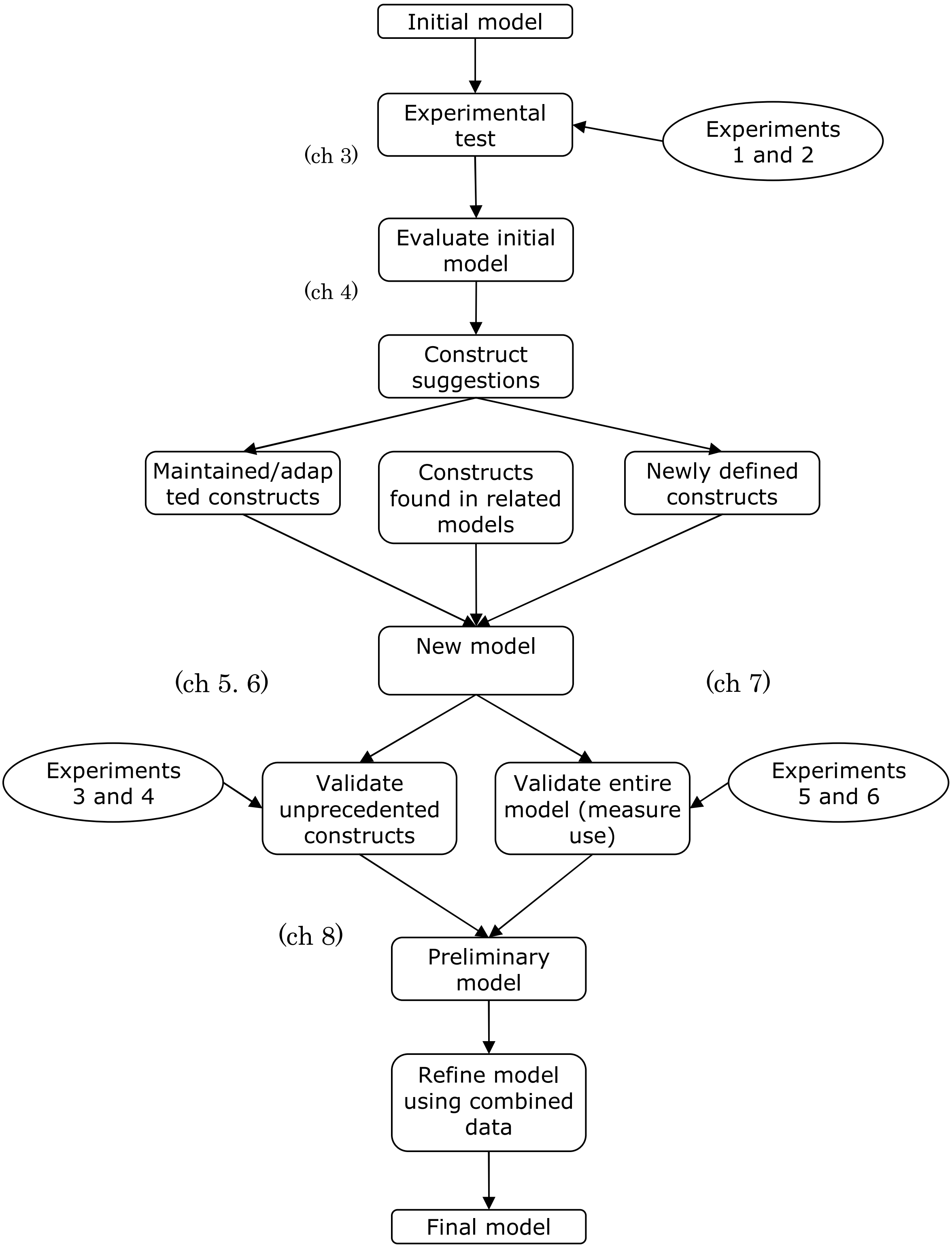


Figure 4.1. Strategy visualization

3. Design the model by combining the found constructs and the established hypothetical interrelations between these constructs.
4. Develop measurement instruments and refine them after testing. The main instrument is, as before, a questionnaire with statements that can be replied to on a Likert type scale, but additionally we can use the results of an observation analysis.
5. Validate the model in two steps:
   a. - Establish the relevance and validity of constructs that have not been included in any acceptance model yet.
   b. - Test the model in experimental settings with different types of assistive social robots and validate it with usage data (Intention to Use should predict actual use of the system).
6. After these tests, the model can be refined by confirming or rejecting hypothetical connections between the constructs. In this refining stage, we will combine the data that has been collected in the different experiments and apply structural equation modeling.

We will discuss items 1 to 4 in the following sections of this chapter. Item 5 (model validation) will be subject to Chapters 5 to 7 and the refining of the model (item 6) will be subject to Chapter 8.

There are also two additional issues that will be addressed alongside the model development. First there is behavior analysis, a technique that provided us with useful insights after the first experiments. We will refine the methodology in Chapter 5. Second, there is the issue of moderating influences: the strength of the influence of constructs upon each other is found to be relative to user characteristics like age, gender, experience and education (Venkatesh et al. 2003; Sun and Zhang 2006). We will go into further detail about this issue in Section 4.4.4.

4.3 A further evaluation of experiments 1 and 2

4.3.1 Implications of previous analyses

In Chapter 3 we found that Attitude was a construct with a strong influence on Intention to Use when added to the determining constructs in a regression analysis (see Table 3.16). We concluded it thus had to be included in a new model as a direct determinant. This is in concordance with the findings of related studies as presented in section 2.4.3.

Also Conversational Acceptance and Social Abilities were found to be of influence (see also Table 3.16) when added in the regression analysis on the results for Annie. These are ‘unprecedented constructs’: they have never been part of an acceptance model. Nevertheless, the underlying concept of feeling comfortable with a robot as a conversational partner should be represented in a new model, as well as the concept of Social Abilities. These two concepts may even be represented within one construct, since accepting someone as a conversational partner can be part of social acceptance.

We thus identify Attitude, Conversational Acceptance and Social Abilities as concepts that are relevant to the new model and will adopt them as constructs.

4.3.2 Construct suggestions from principal component analysis

An indication for alternative constructs could come from a principal component analysis on the results of the experiment as described in the previous chapter. This method is often used to detect hidden factors which underlie the detected relationships among the questions beyond the existing constructs (Malhotra and Galletta 1999; Mathieson et al. 2001; Shibata et al. 2003). These hidden factors could either indicate the possibility of new constructs or suggest a relation between constructs that has not been suspected before.

	Component				
	1	2	3	4	5
3 CA	-.014	-.083	.245	**.645**	.318
4 PE	.276	**.782**	.106	.117	-.007
5 PE	.269	**.781**	.039	.188	.165
6 EE	.490	.237	**.498**	.350	-.061
7 EE	**.782**	.123	.167	.327	.109
8 EE	**.629**	.360	.215	.047	.120
9 SI	.056	.308	-.140	**.667**	.139
12 SI	.116	**.704**	-.095	.008	.163
13 SA	.322	.148	.506	**.629**	-.096
14 SA	.127	.095	**.720**	.190	.050
15 SA	.156	.304	.101	.138	**.846**
16 SA	.226	.115	.183	.197	**.851**
17 SA	.000	.015	**.759**	-.058	.263
18 AT	.217	**.726**	.458	.032	.132
19 AT	.387	**.581**	.290	.324	.106
20 SE	**.770**	.230	-.038	.133	.093
21 SE	**.750**	.139	.032	.194	.022
22 SE	**.798**	.263	.033	-.116	.133
23 ANX	.484	.300	.293	**.534**	.106
24 ANX	**.481**	.081	.182	-.252	.221
Eigenvalues	4.06	3.26	2.23	2.16	1.84
% of variance	20.32	16.30	11.17	10.79	9.21

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization (Rotation converged in six iterations).
Note: the highest score for each item appears bold.

Table 4.1. Principal component analysis - rotation component matrix using the combined scores of both robots

We applied a principal component analysis with varimax rotation for this, using the combined scores of the agents. We added the questions on Social Abilities and Conversational Acceptance, and omitted questions 10 and 11 from the results, because these were dependent questions (depending on question 9) in the SI-construct. The questions on Intention to Use were also not included, because of their unique interrelation (they ask the same question, just the time is varied) and can thus form a construct that other constructs are to relate to.

In order to determine the components, we used Kaiser's criterion, which limits the outcome to those with initial eigenvalues over 1.0 (Kaiser 1960; Kaiser 1970). The rotation converged in seven iterations and identified five distinct components. Table 4.1 shows the results.

If we look at the construct items and their division among the components, we see that the items of the constructs EE, SE and ANX score high on the first component. However, for EE item 6 scores (slightly) higher on the third component and for ANX item 23 scores higher on the fourth component. Just for the constructs PE, AT and SE, all items that form each construct have their highest score on the same component (2 and 1).

Based on these results, we distinguished five factors on which the responses to the questionnaire items loaded. We regrouped the items accordingly (see Table 4.2), thus forming the following 'component based constructs':

- The first factor we called Effort, Ease and Anxiety (EEA). It measures the extent to which people think they can adapt easily to the technology, learn how to work with the technology and overcome eventual anxieties. It features the questionnaire items 7, 8, 20, 21, 22 and 24. We could have included items 6 and 23, because they score high on component 1 and would thus be united with the items that belonged to their original constructs (EE and ANX). However, we included them in the third and fourth construct because (1) they scored slightly higher on the related components, (2) the first construct already contained six items and (3) our focus is on exploring new constructs rather than confirming the existing ones.
- The second factor we called Performance and Attitude (PA). It measures performance related attitude towards the usefulness of the new technology. It features the questionnaire items 4, 5, 12, 18 and 19.
- The third factor is called Sociability (SO). It measures how sociable the participants rate the system and includes how well they feel understood. It features questionnaire items 6, 14 and 17.
- The fourth factor we called Perceived Enjoyment (PE). It incorporates questionnaire items 3, 9, 13 and 23, and measures how pleased people feel and how pleased they think others will feel with the new technology. This construct contains questions that originally addressed various concepts (i.e. social influence, anxiety), but in this particular context are related to enjoyment.
- The fifth factor we called Trust (TR). It features questionnaire items 15 and 16 on how well the subjects trust the system when it advises them.

Factor	#	Questions	Originally
Effort, Ease and Anxiety (EEA)	*1*	*Do you think you can quickly learn how to control the robot?*	*7 (EE)*
		Do you think the robot is easy to use?	*8 (EE)*
		Do you think you could work with the robot without any help?	*20 (SE)*
		Do you think you could work with the robot if you could call someone for help?	*21 (SE)*
		Do you think you could work with the robot if you had a good manual?	*22 (SE)*
		If you were to use the robot, would you be afraid to make mistakes or break something?	*24 (ANX)*
Performance and Attitude (PA)	*2*	*Do you think the robot would be useful to you?*	*4 (PE)*
		Do you think the robot would help you do things?	*5 (PE)*
		Do you think the staff would be pleased if you would have this robot?	*12 (SI)*
		Do you think it is a good idea to use the robot?	*18 (AT)*
		Would you like to use the robot?	*19 (AT)*
Sociability (SO)	*3*	*As you have noticed, you control the robot by speech. Do you think you can easily communicate with it that way?*	*6 (EE)*
		Would you consider the robot to be social?	*14 (SA)*
		Do you feel understood by the robot?	*17 (SA)*
Perceived Enjoyment (PE)	*4*	*Did you feel uncomfortable talking to a robot?*	*3 (CA)*
		Do you think many people would be pleased if you would have this robot?	*9 (SI)*
		Did you find the robot a pleasant conversational partner?	*13 (SA)*
		Do you feel at ease with the robot?	*23 (ANX)*
Trust (TR)	*5*	*Would you trust the robot if it gave you advice?*	*15 (SA)*
		Would you follow the robot's advice?	*16 (SA)*

Table 4.2. Constructs based on the principal component analysis

4.3.3 Additional calculations and conclusions

To establish whether the constructs based on the components would form a model that performed better than the UTAUT model in Chapter 3 on robustness and explanatory power, we performed additional calculations.

First, the Cronbach's Alpha calculation on the results of the test of our initial UTAUT model in the previous chapter (see Table 3.14) showed two constructs that had a low score: Social Influence and Anxiety. In the above section we saw how the items of these constructs indeed had a high score on different components. When we applied Cronbach's Alpha to the component based constructs (Table 4.3), we found that overall they appear more consistent than the UTAUT constructs, although the construct of Sociability does not pass the .7 threshold. If one of the three items is be omitted form this calculation, it is even lower (.573, .459 and .517). If we look at the original construct of Social Abilities (Table 3.4), we find this to be reliable. Apparently, the reduction of the number of items has resulted in a less reliable construct.

Secondly, we addressed the issue of the effect of the more and less social condition: in the previous chapter we also saw that there was an effect of the increased sociability on user behavior, but this was not reflected in the UTAUT constructs. When we performed a Mann-Whitney U-test for the component based constructs, comparing the more and less social conditions, we see that there is a small, but clear effect (Table 4.3). The scores are generally higher for the more social condition and there is a significant difference on Perceived Enjoyment. Apparently, a more social robot is perceived to be more enjoyable.

Component	Cronbach's Alpha	Mann-Whitney U	Sig. (2-tailed)
Performance and Attitude	.861	580	0.446
Effort Ease and Anxiety	.848	620	0.759
Sociability	.619†		
Trust	.886	662	0.083
Perceived Enjoyment	.703	597	0.034*
Intention to Use	.875	556	0.296

†Internal consistency for construct below threshold.

Table 4.3. Cronbach's alpha for the alternative constructs and U-scores for the two conditions

A third additional calculation concerned a regression analysis, using the component based constructs as determining influences on Intention to Use. Table 4.4. presents the results, showing:

- Performance and Attitude as the dominating determining influence on Intention to Use for both agents and both conditions;
- a significant influence of Perceived Enjoyment for Annie;
- a high R^2 score, especially for Annie;
- a difference of .20 in the R^2 score for the two agents – this is less than the difference between the agents for the original UTAUT constructs (Table 3.14) and the agents for UTAUT with added constructs (Table 3.18);
- a difference of .02 in the R^2 scores for the two conditions.

We conclude the following from this regrouping effort:

- If we use the component based constructs, the model performs almost equally well in two conditions, but (still) not equally well for the two agents.
- Performance and Attitude forms a reliable construct, implying that the original UTAUT constructs Perceived Usefulness and Attitude are very close. It should be examined (also by looking at related research) whether Attitude can be seen as a determining influence on Perceived Usefulness.
- For Effort, Ease and Anxiety we have a similar conclusion: the original UTAUT constructs Perceived Ease of Use and Anxiety are very close and we should examine interdependence.
- The construct of Trust could indeed be added, although its interdependence with Social Abilities should be examined, since the questions of this construct were part of the original (reliable) construct of Social Abilities.

Agent type	*Independent*	*Dependent*	*Beta*	*t*	*Sig.*	R^2
iCat	*Performance and Attitude*	*Intention to*	*.714*	*4.045*	*.000***	*.59*
	Effort Ease and Anxiety	*Use*	*.197*	*1.161*	*.255*	
	Perceived Enjoyment		*-.220*	*-1.493*	*.145*	
	Trust		*.038*	*.303*	*.764*	
Annie	*Performance and Attitude*	*Intention to*	*.662*	*6.130*	*.000***	*.79*
	Effort Ease and Anxiety	*Use*	*.093*	*.877*	*.387*	
	Perceived Enjoyment		*.257*	*2.432*	*.021**	
	Trust		*.008*	*.081*	*.936*	
More social	*Performance and Attitude*	*Intention to*	*.569*	*4.090*	*.000***	*.64*
	Effort Ease and Anxiety	*Use*	*.150*	*1.022*	*.316*	
	Perceived Enjoyment		*.187*	*1.361*	*.184*	
	Trust		*.096*	*.771*	*.447*	
Less Social	*Performance and Attitude*	*Intention to*	*.792*	*4.319*	*.000***	*.66*
	Effort Ease and Anxiety	*Use*	*.147*	*.911*	*.369*	
	Perceived Enjoyment		*-.128*	*-.850*	*.402*	
	Trust		*-.018*	*-.139*	*.890*	
Combined	*Performance and Attitude*	*Intention to*	*.671*	*6.790*	*.000***	*.65*
	Effort Ease and Anxiety	*Use*	*.138*	*1.470*	*.146*	
	Perceived Enjoyment		*.067*	*.740*	*.462*	
	Trust		*.011*	*.128*	*.899*	

Table 4.4. Regression scores for component based constructs

- The newly formed construct of Sociability does not result in a reliable construct, while the original construct of Social Abilities did (see Table 3.4). It would thus be a good idea to further adapt the construct.
- The items in the construct of Social Influence should be rephrased: apparently they are interpreted in a way that associates them with enjoyment or performance/attitude since they score high on components that we identified as such – which also explains the low Cronbach's Alpha in Table 3.4.
- The construct of Perceived Enjoyment should be added. Its definitive form can be established after comparing it to similar constructs in related research.

4.4 Constructs for a new model

We will now identify the constructs that we hypothesize to form an accurate model for measuring acceptance of assistive social agents used by older adults. In order to connect to related research, we will maintain constructs where possible, but revise their contents and interdependences as suggested by the conclusions in the previous section.

First we will look into constructs that will be maintained with only minor adaptations. Next we will look into constructs that have been split, combined or altered otherwise and subsequently we will establish constructs that are relevant to our technology and user group according to related research. Finally we will establish which moderating factors could be relevant.

The questionnaire items that we developed for the constructs that are discussed in the following section can be found in Table 4.6.

4.4.1 Maintained constructs

Perceived Ease of Use and Perceived Usefulness

Generally in technology acceptance modeling, when new models are developed or existing models are modified the original TAM constructs *Perceived Ease of Use and Perceived Usefulness* as introduced by Davis (Davis 1989; Davis 1993) are maintained with just minor adaptations in the questionnaire items. In some studies the model was adapted with other minor changes, for example by the introduction of just one other construct, or the way the constructs influence each other, but the two constructs Perceived Ease of Use and Perceived Usefulness are always maintained as determinants of Intention to Use (see overviews in Lee et al. 2003; Benbasat and Barki 2007; Yuanquan et al. 2008). Also in the UTAUT model the original TAM constructs were maintained, although they were defined in a broader sense and named Effort Expectancy and Performance Expectancy (see section 2.4.3).

We thus conclude that Perceived Ease of Use and Perceived Usefulness can be seen as the basic constructs that are shared by all different models. We therefore maintain these constructs and since the results from our previous experiments showed them to be reliable, we can also maintain the used questionnaire items.

Considering the construct interrelations, in technology acceptance models both are generally found to be influencing Intention to Use. However, the two constructs can be interrelating: many studies find Perceived Ease of Use to be a determining influence on Perceived Usefulness (Davis 1993; Lee et al. 2003; Venkatesh et al. 2003; Wilson and Lankton 2004; Yang and Yoo 2004; Sun and Zhang 2006). This means that the tested system is partly perceived as useful because of its easiness of use, which could very well be applicable to an assistive social robot. We will therefore investigate this possible interdependence and note this as a hypothetical influence.

One of our conclusions after the principal component analysis was, that it is very likely that Attitude is a determining influence on Perceived Usefulness. This is not a new construct interrelation, since it has been confirmed by several TAM studies (e.g. Chau 2001; Yang and Yoo 2004).

Also based on the conclusions of our principal component analysis is our decision to expand the construct of Perceived Ease of Use with items that were in the UTAUT model part of the construct of Self Efficacy. This construct turned out not to be reliable for both experiment 1 and 2, while its items combined with items of Perceived Ease of Use in our alternative construct analysis lead to a reliable construct (Table 4.3).

Attitude

As mentioned above, the influence of Attitude can be subject to discussion, for it can be a direct influence to Intention to Use or through Perceived Usefulness. Research by Yang and Yoo (Yang and Yoo 2004) states that Attitude can be interpreted as *affective* attitude (refers to how much the person likes the object of thought) or *cognitive* attitude (refers to an individual's specific beliefs related to the object). According to Yang and Yoo (2004), the first factor only seems a moderating influence on the acceptance process, while the second one is a stronger and more direct influence and should be considered a direct determinant of Intention to Use.

The direct influence of Attitude on Intention to Use is confirmed by related research (Wu and Chen 2005) and by our own findings, both by the regression analysis we performed in section 3.7.3 (see Tables 3.16-3.18) and partly by our conclusions from the principal component analysis (see Table 4.4).

Social Influence

Also the influence of the construct of Social Influence (also called social norm) (Sun and Zhang 2006) can be subject to discussion. In our experiments we found it not to be a reliable construct. As we concluded at the end of section 4.3.3, the regrouping of the constructs based on the principal componant analysis results showed they are interpreted in a way that associates them with either enjoyment or performance/attitude. This means that if we maintain the construct, the items should be rephrased.

Considering the interrelations: studies that include this construct find it to have a determining influence both directly on usage and on Intention to Use (Malhotra and Galletta 1999; McFarland and Hamilton 2006). We will test these influences within our new model.

Anxiety

For the combined factors within our second construct, called Effort, Ease and Anxiety (EEA) we did not find support in related literature. On the contrary: anxiety is subject to many other robot acceptance related studies that do show its relevance in interaction with robots (Nomura et al. 2006) and screen agents (Cowell and Stanney 2005), but these studies do not have the goal to model acceptance. There are acceptance model studies that successfully claim its determining influence on Perceived Ease of Use and Perceived Usefulness, but they concern the acceptance of technology that is very different from robots, like software packages (Montazemi et al. 1996) and information or communication technology (Gopal et al. 1997; Schaper and Pervan 2007).

We thus want to keep Anxiety as a separate construct, but hypothesize its interdependence with Perceived Ease of Use and Perceived Usefulness. Regarding the items, we needed to look for a more reliable set, for the Cronbach's Alpha score of the original construct was far below the .7 threshold. Thus, we omitted the item (23) that was apparently more associated with enjoyment and added items that were derived from the above mentioned anxiety related research on robots and screen agents (Nomura et al. 2006).

Facilitating conditions

The construct of facilitating conditions was not used in our model, because it was not relevant to our setting. It is only relevant in a context where subjects believe the technology to be actually available to them and this was not the case in our experiments. If we are going to validate our model in new experiments however, we have to create a situation in which the tested technology is also actually available for use after the first impression (in terms of our model: the first impression leads to Intention to Use and the validation demands a situation of usage).

The items used for this construct were adapted from the original UTAUT questionnaire.

4.4.2 New constructs

The constructs in this section were formed out of the regrouped questions. So their items (questions) have been used in our previous experiments, but as constructs within our developed model, they are new.

Trust

Our construct of Trust as we formed it from UTAUT items after the principal component analysis has been proven a reliable one (see Table 4.1). In addition it is found to be a relevant influence in other studies on acceptance in general (Marsh et al. 2004; Wu and Chen 2005; Cody-Allen and Kishore 2006) and on human robot interaction (Shinozawa et al. 2003; De Ruyter et al. 2005; Shinozawa et al. 2005).

In the first group of studies, it is claimed to have a direct influence on Intention to Use but in both types of studies, it is also related to either social abilities or social behavior: a robot or screen agent with more social abilities is found to gain more trust by its users. We therefore will explore the interdependence between Trust and Social Abilities and the determining influence of Trust on Intention to Use. To incorporate Trust into the model, two items were developed to measure the trust the user has in the robot and the extent to which the user intends to comply to the robot's advice.

Perceived Enjoyment

Several acceptance studies confirm our conclusion after the principal component analysis that Perceived Enjoyment can be a reliable construct that directly determines Intention to Use (Van der Heijden 2004; Chesney 2006; Sun and Zhang 2006). In evaluating social agents such as companion type robots, an element of pleasure when interacting with the agent may indeed influence user acceptance. Van der Heijden (2004) points out that in ‘hedonic systems’ (system that are mainly used for entertainment), the concept of enjoyment is a crucial determinant for the Intention to Use for these systems. Of course, social agents in eldercare will hardly be developed just to entertain: they will be partly utilitarian, partly hedonic. But even if just partly hedonic, enjoyment is found to be a construct that needs to be part of an acceptance model for robotic technology as is illustrated by a study concerning the Lego Mindstorms development environment by Mindstorms hobbyists (Chesney 2006). The study, based on the viewpoint that this concerns a partly hedonic, partly utilitarian type of system, confirms Perceived Enjoyment having just an indirect effect on Intention to Use.

But even in utilitarian systems Perceived Enjoyment can be a relevant influencing factor, as pointed out in an extensive study by Sun and Zhang (2006). The study mainly supports the claims by Venkatesh et al. (Venkatesh 2000) and Yi and Hwang (Yi and Hwang 2003), that Perceived Enjoyment has no direct influence on Intention to Use, but that it can influence Ease of Use and Usefulness. Still the study does also recognize that this is not a general claim for all types of systems. Indeed this could work very differently for robotic systems used by elderly people.

We conclude that both the above mentioned acceptance studies and our own findings justify the introduction of Perceived Enjoyment in our model. The perceived enjoyment items we used for our questionnaire (section 4.6) were adapted from Davis et al. (1992).

Perceived Sociability

Our construct of Perceived Sociability is unique as a part of an acceptance model. This is not surprising, since there has been no other attempt to create an acceptance model for our specific field of autonomous interactive systems. Nevertheless, the need for social abilities for robots to function effectively as assistive devices has been established in earlier studies (see section 2.2), and it is not surprising that these abilities are found to be of influence on the appreciation of robots in general (Forlizzi 2007; Mitsunaga et al. 2008). In addition, the added construct of social abilities in our first experiment correlated with all UTAUT constructs (see Table 3.9), demonstrating that it relates to aspects of acceptance, although its interdependence with these constructs still has to be established.

Since what we measured is not actual social abilities, but the degree to which the user perceives the robot to be sociable, we decided to rename the construct Perceived Sociability. The item on Conversational Acceptance was also integrated.

As stated earlier, we suspect that Trust has a determining influence on Perceived Sociability, based on claims and findings found in related work.

4.4.3 Additional constructs

The following constructs have been formed on the basis of observations of elderly users in their interaction with a robot or screen agent during our experiments. They have not been a part of any acceptance model yet, but as we will describe, we found clues on their relevance in this context in several related studies. Furthermore they are – like the constructs of Perceived Sociability and Trust - linked to the notion of social acceptance.

Perceived adaptivity

When presenting a robot to elderly users in our previous experiments, we told these users the possibilities: it could help them remember things, it could help them control all kinds of devices, it could watch them and alarm someone if necessary and it could keep them company, play games or just chat. During and after their encounter with the robot, a repeating remark was that they would not have any use for it, because they could still remember a lot. Therefore, when questioned if they would intend to use the robot if it were available to them, they would reply negatively. When we suggested that the features they would not yet need could simply be turned off by them, or automatically by the robot itself if it would notice that there was no demand for it, the participant would change her mind and be willing to at least try to use the robot. Most of these subjects were suffering light forms of dementia.

This illustrates that the extent to which a system is perceived to be adaptive to the different stages of aging could very well be of influence on acceptance of it. This has to do with the fact that the conditions of elderly people change over time. Mobility may improve after a hip replacement, heart condition may deteriorate or improve after changes in medication, eyesight, hearing and dementia may become worse over time. Requirements for the type of support that is needed, therefore change over time. Intelligent assistive technology would need to adapt to these changes in conditions in order to provide appropriate support. Previous studies that address elderly users argue that Adaptivity is an essential aspect of technology that is developed for aging users (for an extensive overview of adaptivity and aging in general see Pew and Hemel 2004). But there is also more specific research concerning robot and screen agent technology. Research by Maciuszek and Shahmehri (Shahmehri 2001; Maciuszek and Shahmehri 2003) for example, focuses on setting the specifications for 'multifunctional adaptive virtual companions for later life' with arguments based on requirements research. Moreover, Forlizzi et al. included adaptivity in their design guidelines for robotic products that 'support the ecology of aging', based on long term qualitative research, observing the changing needs of a group of Midwestern aging people, learning how these products might assist these people, helping them to stay independent and active longer (Forlizzi et al. 2004). These studies suggest that when users perceive the system to adapt to their changing needs, they will find it more useful and will be more accepting towards the system.

However, this literature is not decisive on whether this concerns adaptability (users being able to adapt the system) or adaptiveness (the system adapting by itself). Although we can form questionnaire items that ignore this difference, we still need to find out which of those is preferred by elderly users. Furthermore, we also need to investigate if this construct would be of direct influence on Intention to Use or an influence on Perceived Usefulness (the more adaptive, the more useful the system is).

In chapter 6 we will further discuss the impact and interpretation of this construct.

Social Presence

Since it is not unusual for humans to engage with technology as if it were a social entity (Reeves and Nass 1996) it can be expected that this effect is exacerbated when technology takes the form of an embodied character and interacts in a social manner using natural language and non-verbal human behaviors. The extent to which embodied systems are engaged as social entities appears to be related to a factor that is often related to as either 'Presence' or, more specifically 'Social presence' (DiSalvo et al. 2002; Lee and Nass 2003; Bickmore and Schulman 2006).

The term presence originally refers to two different phenomena. First, it relates to the feeling of really being present in a virtual environment and can be defined as 'the sense of being there' (Witmer and Singer 1998). Second, it can relate to the feeling of being in the company of a social entity: 'the perceptual illusion of nonmediation' (Lombard and Ditton 1997). In our context, the second definition, interpreted as 'the sense of a social entity communicating directly with the user' (Biocca et al. 2003) is relevant, since the extent to which one feels to be dealing with a social entity when meeting a robot or screen agent is of influence on the way it is perceived and accepted.

This is found in studies by Lee and Nass and Bickmore et al. (Lee and Nass 2003; Bickmore 2004; Bickmore et al. 2005; Bickmore and Schulman 2006; Bickmore and Schulman 2007; Bickmore et al. 2008), in which it is also established that social presence is related to the sociability of the system. In the case of Lee and Nass the 'social entity' only manifested itself vocally, demonstrating a more extraverted voice caused a higher sense of presence. In the study by Bickmore et al. a relational agent is developed and the possibility of social bonding is explored. It is found that social abilities do influence presence and that this leads to enjoyment experienced by the users.

So we intend to incorporate measuring social presence in our experiments to research its role and establish its relationship with social abilities and perceived enjoyment. We expect the following, based on the above mentioned studies:

- the sense of presence increases if a system is perceived to have more social abilities;

- Perceived Enjoyment increases if sociability is perceived stronger and the sense of social presence is felt more intensely.

To include this construct in the model, a set of 5 items were developed, adapted from Bailenson et al. (2001).

4.4.4 Moderating factors

As we stated earlier in Section 4.2, the strength of the influence of constructs upon each other is found to be relative to user characteristics. These moderating factors are influences that are no aspects of individual subjective judgment, but that do influence this judgment or the way it leads to Intention to Use or usage. In acceptance methodology, they are a standard part of a model, either as general influences on all processes within the model or with specific influences on certain interdependences (Sun and Zhang 2006).

The UTAUT model states that *Age*, *Gender* and *Experience with technology* can be moderating factors. However, there are other factors that could influence the process of acceptance – as listed by Sun and Zhang – that can be divided into the categories of:

- organizational factors (*voluntariness, tasks and perhaps facilitating conditions*),
- technology related factors (*purpose, complexity, perhaps social abilities, and perhaps adaptability*) and
- individual factors (*intellectual capability, cultural background, gender, age, affective attitude, experience, physical condition*).

In general, these factors are related to a technology type and a user group that are both very different from ours. Organizational factors are not relevant, because our participants are not part of an organization. Technology related factors that are relevant are already represented in proposed constructs. Individual factors gender, age and experience could be relevant. In our case, where hardly any participant will have worked with a robot, we think experience could be broadened to the less specific computer experience. In addition, we are interested in education level as an individual factor, since this might be related to anxiety and attitude (Burton-Jones and Hubona 2005; Gumussoy et al. 2007).

Thus we plan to collect data on gender, age, education level and computer experience when using the questionnaire and will analyze results to see which processes are influenced by these factors.

Based on the above mentioned literature, we expect the influence of the factors to be as follows (we will call these assumptions moderating factor hypotheses (MFH)):

MFH1 *The influence of Perceived Usefulness on Intention to Use is moderated by (a)Gender and (b)Age.*

MFH2 *The influence of Perceived Ease of Use on Intention to Use is moderated by (a)Gender, (b)Age, and (c)Experience.*

MFH3 *The influence of Social Influence on Intention to Use is moderated by (a)Gender, (b)Age and (c)Experience.*

MFH4 *The influence of Facilitating Conditions on Usage is moderated by (a)Age and (b)Experience.*

MFH5 *The influence of (a)Anxiety on Perceived Usefulness and (b)Perceived Ease of Use will be moderated by Education.*

To establish the influence of these moderating factors within our new model, we will perform a Chow's test. As indicated in Section 2.5, this is a common UNIVARIATE type procedure for establishing the strength of moderating influences (Sharma et al. 1981). We will perform this test on the results of each experiment after a regression analysis has indicated which interrelations are significant. Hypotheses on moderating influences that relate to construct interrelations that are not significant are of course not applicable for the specific data set. If, for example, a regression analysis on the results of an experiment shows Perceived Usefulness as a non significant determining influence on Intention to Use, MFH1 will be not applicable.

4.5 Model overview

In the previous section we mentioned all the constructs that we are going to embed in our model and hypothesized their interdependences. The following list summarizes these hypotheses.

As both the model overview and the constructs overview (see Figure 4.2. and Table 4.4.) show, the model incorporates the following hypothetical connections:

H1 *Use is determined by (a) Intention to Use, (b)Social Influence and (c)Facilitating Conditions.*

H2 *Intention to Use is determined by (a) Perceived Usefulness, (b) Perceived Ease of Use, (c) Attitude, (d) Perceived Enjoyment, (e) Social Influence and (f) Trust.*

H3 *Perceived Usefulness is determined by (a) Anxiety, (b) Attitude, (c) Perceived Adaptivity and (d) Perceived Ease of Use*

H4 *Perceived Ease of Use is determined by (a) Anxiety and (b) Perceived Enjoyment*

H5 *Perceived Enjoyment is determined by (a) Social Presence and (b) Perceived Sociability*

H6 *Perceived Sociability is determined by Trust*

H7 *Social Presence is determined by Perceived Sociability*

Figure 4.2 vizualizes these hypothetically interrelations.

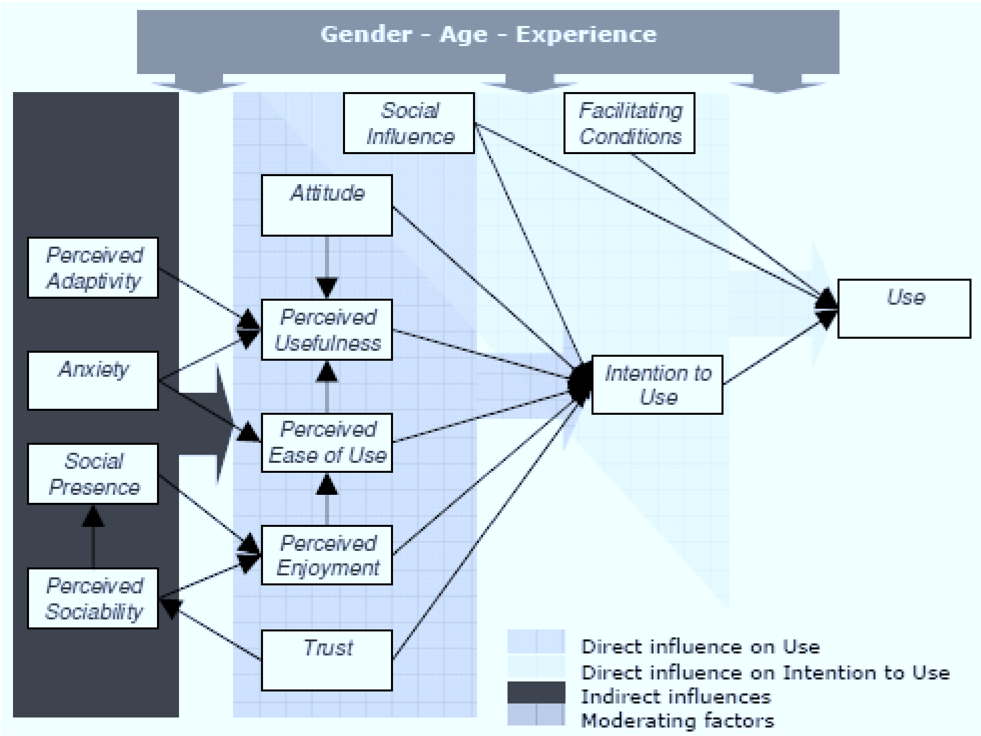


Figure 4.2. Model overview

For these hypotheses we found the following support (see section 4.4):

- H1(a) is supported by (Davis 1986)(basic TAM assumption)
- H1(b) is supported by and (Malhotra and Galletta 1999)
- H1(c) is supported by (Venkatesh and Davis 2000) and (Venkatesh et al. 2003)
- H2(a) and H2(b) are supported by (Davis 1989) and (Venkatesh et al. 2003)
- H2(c) is supported by (Yang and Yoo 2004), (Wu and Chen 2005), (Pynoo et al. 2007) and (Knutsen 2005)
- H2(d) is supported by (Chesney 2006), (Van der Heijden 2004), (Venkatesh and Davis 2000) and (Sun and Zhang 2006)
- H2(e) is supported by (Venkatesh et al. 2003)
- H2(f) is supported by (Lee and Nass 2003) and (Shinozawa et al. 2005)
- H3(a) is supported by (Montazemi et al. 1996) and (Gopal et al. 1997)
- H3(b) is supported by (Chau 2001) and (Yang and Yoo 2004).
- H3(c) is supported by arguments found in (Forlizzi et al. 2004), (Camarinha-Matos and Afsarmanesh 2002), (Pew and Hemel 2004) and by (Maciuszek and Shahmehri 2003)
- H3(d) is supported by (Davis 1986)
- H4(a) is supported by (Schaper and Pervan 2007), (Montazemi et al. 1996) and (Gopal et al. 1997)
- H4(b) is supported by (Yi and Hwang 2003), (Venkatesh 2000) and (Sun and Zhang 2006)

- H5 is supported by(Bickmore and Schulman 2006) and (Lee and Nass 2003)
- H5(b) is supported by (Lee and Nass 2003) and (Bickmore and Schulman 2006)
- H6 is supported by arguments found in (Shinozawa et al. 2005) and (Mitsunaga et al. 2008)
- For H7 we found support in (Shinozawa et al. 2003)

Table 4.5 lists all the constructs and their interrelations, and it gives a short definition for each construct.

4.6 Instruments: questionnaire and user observation

We developed a questionnaire in which each construct is represented by multiple items that could be replied to on a (five point) Likert scale.

Code	***Construct***	***Definition***	***Determined by***	***Determining***
ANX	*Anxiety*	*Evoking anxious or emotional reactions when using the system.*		*PU, PEOU*
ATT	*Attitude*	*Positive or negative feelings about the appliance of the technology.*		*ITU*
FC	*Facilitating conditions*	*Objective factors in the environment that facilitate using the system.*		*Use*
ITU	*Intention to Use*	*The outspoken intention to use the system over a longer period.*	*ATT, PU, PEOU, Trust*	*Use*
PAD	*Perceived adaptivity*	*The perceived ability of the system to be adaptive to the changing needs of the user.*		*PU*
PENJ	*Perceived enjoyment*	*Feelings of joy or pleasure associated by the user with the use of the system.*	*SP, PS*	*ITU, PEOU*
PEOU	*Perceived ease of use*	*The degree to which the user believes that using the system would be free of effort*	*ANX, PENJ, PS*	*ITU, PU*
PS	*Perceived sociability*	*The perceived ability of the system to perform sociable behavior.*	*Trust*	*PENJ, SP*
PU	*Perceived usefulness*	*The degree to which a person believes that using the system would enhance his or her daily activities*	*ANX, PAD, PEOU*	*ITU*
SI	*Social influence*	*The user's perception of how people who are important to him think about him using the system*		*ITU, Use*
SP	*Social presence*	*The experience of sensing a social entity when interacting with the system.*	*PS*	*PENJ*
Trust	*Trust*	*The belief that the system performs with personal integrity and reliability.*		*ITU, PS*
Use	*Use/Usage*	*The actual use of the system over a longer period in time*		*ITU, FC, SI*

Table 4.5. Constructs overview

Items that were adapted from UTAUT were based on original constructs, but adapted to the specific context of robots and elderly users. The construction of non UTAUT items is discussed in sections 4.4.2 and 4.4.3.

ANX	*1. If I should use the robot, I would be afraid to make mistakes with it*
	2. If I should use the robot, I would be afraid to break something
	3. I find the robot scary
	4. I find the robot intimidating
ATT	*5. I think it's a good idea to use the robot*
	6. The robot would make my life more interesting
	7. It's good to make use of the robot
FC	*8. I have everything I need to make good use of the robot.*
	9. I know enough of the robot to make good use of it.
	10. There is someone near me who is available to help me if necessary
ITU	*11. I think I'll use the robot during the next few days*
	12. I am certain to use the robot during the next few days
	13. I'm planning to use the robot during the next few days
PAD	*14. I think the robot can be adaptive to what I need*
	15. I think the robot will only do what I need at that particular moment
	16. I think the robot will help me when I consider it to be necessary
PENJ	*17. I enjoy the robot talking to me*
	18. I enjoy doing things with the robot
	19. I find the robot enjoyable
	20. I find the robot fascinating
	21. I find the robot boring
PEOU	*22. I think it will be difficult to let the robot know what I want.*
	23. I think I will know quickly how to use the robot
	24. I find the robot easy to use
	25. I think I can use the robot without any help
	26. I think I can use the robot when there is someone around to help
	27. I think I can use the robot when I have a good manual.
PS	*28. I consider the robot a pleasant conversational partner*
	29. I find the robot pleasant to interact with
	30. I feel the robot understands me.
	31. I think the robot is nice
PU	*32. I think the robot is useful to me*
	33. It would be convenient for me to have the robot
	34. I think the robot can help me with many things
SI	*35. I think the staff would like me using the robot.*
	36. I think many people would like me having the robot.
	37. I think it would give a good impression if I should use the robot.
SP	*38. When interacting with the robot I felt like I'm talking to a real person*
	39. It sometimes felt as if the robot was really looking at me
	40. I can imagine the robot to be a living creature
	41. I often think the robot is not a real person.
	42. Sometimes the robot seems to have real feelings
Trust	*43. I would trust the robot if it gave me advice.*
	44. I would follow the advice the robot gives me.
	45. If I would give the robot information, it would not abuse this.

Table 4.6. Constructs questionnaire items

For user observation we wanted to use a more refined technique in which behaviors were not only counted, but also weighed: the score on each item would include a degree of certainty and a degree of intensity. Thus, to each counted item, the observers attributed two values: one for the strength (weight) of it and one for the certainty of the observer. Both could be one, two or three points. So if the observer would be sure of someone laughing very loud, this would score two times three points. Our methodology for user observation and the relation of its results to the questionnaire results will be further discussed in Chapter 5.

4.7 Experimental testing

We now developed a model in which UTAUT is extended with Perceived Adaptiveness and with constructs related to social acceptance: Trust, Perceived Sociability, Social Presence and Perceived Enjoyment. With these extensions it has become a model that aims to evaluate both social and functional acceptance of assistive social robots, specifically by older adults. It maps the aspects of acceptance by including influences on actual use, Intention to Use and by indirect influences. In addition, it includes statements on the influence of moderating factors on this process.

In the following chapters we describe a series of experiments that we designed to evaluate the different aspects of this new model. First we discuss an experiment with again the more and less social conditions for the iCat, during which we also used behavior observation. This experiment, which will be described in Chapter 5, focuses on the added constructs that are related to social acceptance: Trust, Perceived Sociability, Social Presence and Perceived Enjoyment. Next, we will focus on adaptivity, both to further clarify the interpretation of the construct and to justify its presence within the model. This will be subject of Chapter 6. In chapter 7 we describe two experiments that were designed to establish the relationship between Intention to Use and actual use, thus validating the predictive power of the model. In chapter 8 we will summarize our findings and establish whether a further exploration of the joined results of the experiments leads us to a further refinement of the model.

5. Measuring the influence of social abilities II

Parts of this chapter have been published earlier in (Heerink et al. 2008b), (Heerink et al. 2009c) and (Heerink et al. 2010a)

5.1 Introduction

Our new model contains a few new constructs that have never been incorporated in an acceptance model: Trust, Perceived Sociability, Social Presence and Perceived Adaptivity. The first three (and the fourth in a lesser sense) represent our goal to establish a model that incorporates the social aspects of robot acceptance. We can therefore refer to them as 'the social constructs'.

These social constructs were derived from the assumption that the sociability of a robot has somehow - directly or indirectly - an effect on its acceptability. Although there are strong arguments from several studies for developing sociable robots in this context, it has not been established yet that a more sociable assistive robot is accepted better than a less sociable robot.

In this chapter we want to do just that: demonstrate that a more sociable robot is accepted better than a less sociable condition of the same robot. And in addition, we expect the scores on the social constructs to provide an explanation for this difference, showing an increased perception of Sociability, Social Presence and Trust for a more social robot. This means that we repeat experiment 1, but with a different instrument: this time we have our new model and we will use our new questionnaire to collect data. Furthermore, we will again look at the results of a user behavior analysis to compare and connect these to our questionnaire results.

5.2. Revisiting the hypothesized influence of social abilities

At the end of chapter 3, we concluded that the UTAUT model we used in our first two experiments did not explain variance equally well in both conditions. In particular, for both used systems, the explained variance on Intention to Use was much lower for a more social condition. We now want to establish whether our new model will show more equal outcomes for two conditions than the UTAUT model. When we used UTAUT in our first experiments (as described in chapter 3) there was a difference of .20 in R^2 scores. This means that in a new experiment – which will be experiment 3 - we want to test the following hypothesis concerning the strength of our model (we add a code for the experiment to avoid confusion with the model validation hypotheses that have been established in Section 4.5):

Exp3 H1 *The difference in explanatory power of the model as expressed by R^2 for the more and a less social condition of the systems will be less than the difference measured for the UTAUT model in experiments 1 and 2.*

Furthermore, we want to establish if the model can be used to research the influence of social abilities by comparing these two conditions. If users are exposed to a more social robot, it can be expected of them to attribute more social abilities to it and experience more of a social entity than when exposed to a less social robot. In terms of our model, we hypothesize that this difference will be reflected in a difference in Perceived Sociability (PS) and Social Presence (SP).

Following the construct interdependences of our model, we expect Perceived Sociability to influence Social Presence (SP). This perceived Social Presence is expected to influence Perceived Enjoyment (PENJ) which can be a direct influence on the Intention to Use the system (ITU) or through Perceived Ease of Use (PEOU).

The constructs involved in this evaluation of the effect of social abilitites concern only a part of our model. We thus have a focus within the model (represented graphically in Figure 5.1) based on the following hypotheses to explain possible differences between a more and less social condition:

Exp3 H2 *The implementation of more social abilities leads to a higher score for Perceived Sociability.*
Exp3 H3 *The implementation of more social abilities leads to a higher score for Social Presence.*
Exp3 H4 *Perceived Sociability is a determining influence on Social Presence*
Exp3 H5 *Social Presence is a determining influence on Perceived Enjoyment.*
Exp3 H6 *Perceived Enjoyment is a determining influence on Perceived Ease of Use*
Exp3 H7 *Perceived Enjoyment and Perceived Ease of Use are determining Intention to Use*

Hypotheses 4 to 7 would be a confirmation of interdependences established in the previous chapter.

5.3 Experiment

As in our first iCat experiment, this setup was designed to be able to compare a robot with more sociability to a less sociable one.

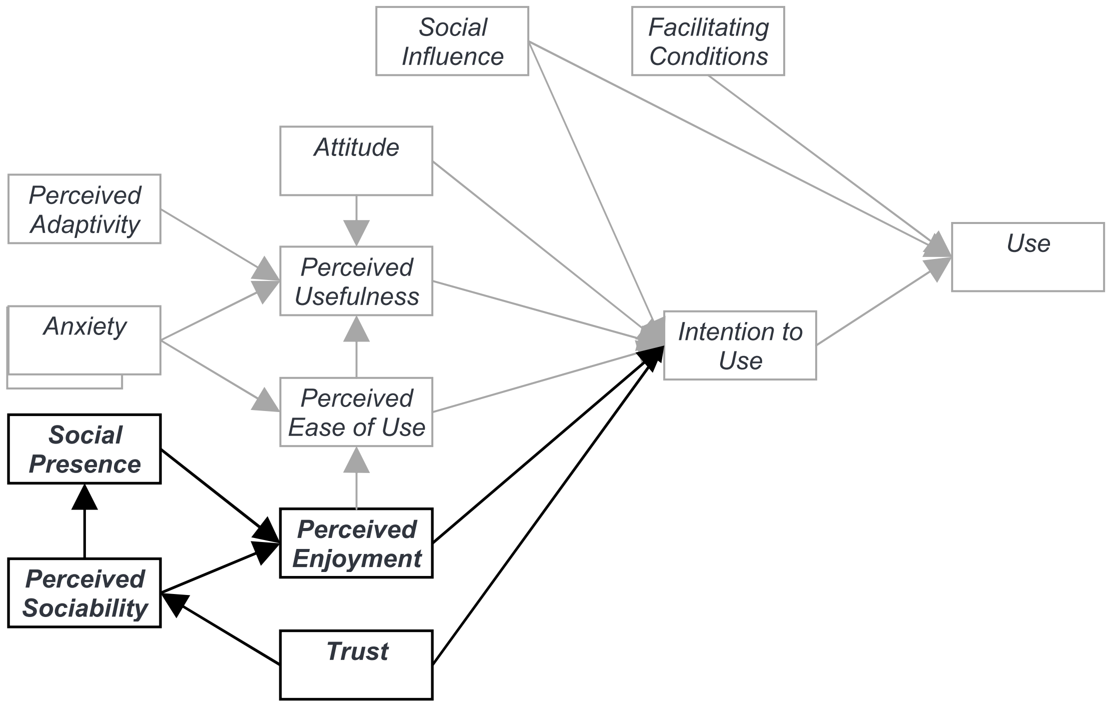


Fig. 5.1. The influence of social abilities

5.3.1 Method

The more socially communicative condition exhibited the same social abilities as in our first experiment. However, the expressiveness in the more social condition was stronger, since it included a more expressive voice by pitch alteration, generated by a more advanced version of our text to speech generator (Loquendo) than the one used in the earlier experiments.

Subjects

Our experiment featured 40 participants between 65 and 89 years old, all living in an eldercare institution in the Dutch city of Lelystad. Exactly half of the participants were exposed to the more sociable version and the other half to the less sociable one. We did not have to omit any participants from the data.

Figure 5.2. Setup iCat Wizard of Oz revisited

Table 5.1 shows the education, computer experience and age for the participants. For Education we used a scale from 1 to 10 – a score of 1 means no education and 10 means university.

	Minimum	Maximum	Mean	Std. Deviation
Age	65	94	82.60	6.291
Experience	1.00	5.00	1.5625	1.105
Education	1	9	3.78	2.178

Table 5.1 Age, computer experience and education of participants (n=40)

Procedure

As in our first experiment, participants were first exposed to the iCat and instructed about its possibilities in groups (eight or four participants per group). After this group session, the participants were invited one by one to have a conversation with the robot, while the other group members were waiting in a different room. Each participant was taken to a separate room to fill out the questionnaire immediately after the individual session. As in the previous experiment the robot's possibilities were: an interface to domestic applications, monitoring, companionship, information providing, agenda-keeping and memorizing medication times and dates. However for this experiment, the robot was only programmed to perform three tasks: setting an alarm, giving directions to the nearest supermarket and giving the weather forecast for tomorrow.

Instruments

We used both the questionnaire based on our model (Table 4.6) and user observation, which we will discuss in detail in Section 5.4.

5.3.2 Model test results

Our calculation of Cronbach's Alpha in Table 5.2 shows the constructs turn out to be reliable. However, to obtain these scores we had to omit statements 10 (FC3), 22 (PEOU1), 37 (SI3) and 45 (Trust3). If we need to omit the same questions for other experiments, we will omit them permanently from our list.

Construct	*Alpha*	*Construct*	*Alpha*
ANX	*.716*	*PEOU*	*.765*
ATT	*.790*	*PS*	*.885*
FC	*.846*	*PU*	*.865*
ITU	*.901*	*SI*	*.752*
PAD	*.740*	*SP*	*.831*
PENJ	*.846*	*Trust*	*.820*

Table 5.2. Cronbach's Alpha

Table 5.3 features the descriptive statistics of the scores. As in our previous experiments, when analyzing the Likert scale replies to the statements, we attributed scores from 1 to 5 to the answers, where 5 would be the most positive score. For 'negative' statements, like those belonging to the construct of Anxiety, we reversed the score so that a higher score still indicated a more positive answer. The statement 'I find the robot scary' was for example noted as a negative statement: the higher the original score, the more scary the robot was experienced. However, in the reversed scores, as presented here, a higher score indicated the experience of a less scary robot.

	Minimum	*Maximum*	*Mean*	*Std. Dev.*
Anxiety	*1.00*	*5.00*	*2.341*	*1.053*
Attitude	*1.00*	*5.00*	*3.383*	*1.023*
Facilitating Conditions	*1.00*	*5.00*	*1.563*	*1.105*
Intention to Use	*1.00*	*5.00*	*3.288*	*1.018*
Perceived Adaptivity	*1.00*	*4.33*	*2.842*	*.955*
Perceived Enjoyment	*2.00*	*5.00*	*3.900*	*.663*
Perceived Ease of Use	*1.00*	*5.00*	*3.525*	*1.054*
Perceived Sociability	*2.00*	*5.00*	*3.750*	*.737*
Perceived Usefulness	*1.00*	*5.00*	*3.200*	*1.088*
Social influence	*1.00*	*5.00*	*3.206*	*.882*
Social Presence	*1.50*	*5.00*	*3.600*	*.914*
Trust	*1.20*	*4.20*	*2.630*	*.879*

Table 5.3. Descriptive Statistics

Table 5.4 shows the t-test scores on the constructs for the two conditions. A positive value for t means the more social condition scores higher than the less social one, a negative number means the opposite. There is a significant difference in acceptance (Intention to Use) score in favor of the more social condition. Also the scores for Social Presence and Perceived Sociability show a significant difference, thus conforming focus hypotheses 2 and 3. Of the social constructs only Trust does not have a higher score for the more social condition.

Construct	**t**	**Mean difference**	**P**
Anxiety	.148	.050	.883
Attitude	-.717	.233	.478
Facilitating Conditions	-.230	.075	.819
Intention to Use	2.264*	.650	.029
Perceived Adaptivity	.000	.000	1.000
Perceived Enjoyment	2.027*	.650	.049
Perceived Ease of Use	.855	.200	.398
Perceived Sociability	2.208*	.333	.034
Perceived Usefulness	.968	.587	.339
Social influence	.342	.100	.734
Social Presence	2.271*	.050	.029
Trust	-.143	.600	.887

Table 5.4. T-test results for the two conditions

Tables 5.5 and 5.6 show the results of a regression analysis on the results. This analysis was applied to the full model, which enables us to compare the results to those of the experiments described in the coming chapters.

The first columns in Table 5.5 show the model validation hypotheses of the full model (Section 4.5) and the second column the hypotheses focusing on the influence of social abilities as established in Section 5.2.

These results show that not all of the model validation hypotheses are confirmed. Intention to Use the robot is predicted by Perceived Ease of Use and Perceived Enjoyment. It may be that the focus on a friendly 'chit-chat'-like interaction with the iCat influenced this focus on comfort rather than competence related factors.

Model validation hypothesis 2 could therefore only be partially confirmed (model validation hypothesis 1 is not included in this analysis because actual voluntary use of the robot over a longer time period was not tested in this experiment).

Validation Hyp.	Focus Hyp.	Indep.	Dependent	Beta	t	P
H2 (a)		PU	ITU	.097	.633	.531
(b)	Exp3 H7	PEOU		.435	3.619**	.001
(c)		ATT		-.037	-.227	.822
(d)	Exp3 H7	PENJ		.581	5.079**	.000
(e)		SI		-.061	-.465	.645
(f)		Trust		.023	.151	.881
Model: R^2=.70; F=13.110; df=6,33; P=.000						
H3 (a)		ANX	PU	.029	.379	.706
(b)		ATT		.381	4.659**	.000
(c)		PAD		.536	6.089**	.000
(d)		PEOU		.077	.928	.357
Model: R^2=.64; F=15.681; df=4,35; P=.000						
H4 (a)		ANX	PEOU	-.430	-3.259**	.002
(b)	Exp3 H6	PENJ		.375	2.845*	.007
Model: R^2=.29; F=7.803; df=2,37; P=.003						
H5	Exp3 H5	PS	PENJ	.526	3.900**	.000
		SP		.331	2.454*	.019
Model: R^2=.61; F=13.294; df=2,37; P=.000						
H6		Trust	PS	.320	2.083*	.044
Model: R^2=.10; F=4.375; df=1,38; P=.043						
H7	Exp3 H4	PS	SP	.540	3.399**	.002
Model: R^2=.43; F=12.952; df=1,38; P=.000						

Table 5.5. Regression analysis on model validation and focus hypotheses

Furthermore, Perceived Usefulness of the iCat was determined by Perceived Adaptivity of the system. Model validation hypothesis 3 can therefore only be partially accepted and model validation hypotheses 4, 5, 6 and 7 can be fully accepted. In addition, the focus hypotheses 4 to 7 can all be accepted.

Overall, Table 5.5 shows how we confirmed the path of influence of social abilities by demonstrating that a higher Perceived Sociability leads to a higher sense of Social Presence, which again leads to a higher score on Perceived Enjoyment. In its turn, Perceived Enjoyment both directly and indirectly (through Perceived Ease of Use) determines Intention to Use.

Table 5.6 specifies the results of a regression analysis on Intention to Use for both conditions. It shows that there remains just a minor difference (.01) in explanatory power as expressed in the R^2 score between the two conditions. This confirms the first hypothesis of this chapter.

Condition	Independent	Dependent	Beta	t	Sig
More Social	PU	ITU	.713	1.368	.195
	PEOU		.194	.813	.431
	ATT		-.428	-.908	.380
	PENJ		.486	2.193*	.047
	SI		-.453	-1.198	.252
	Trust		.405	1.133	.278
	Model: R^2=.72; F=5.539; ; df=6,13; P=.005				
Less Social	PU	ITU	-.063	-.308	.763
	PEOU		.561	3.356**	.005
	ATT		.211	.751	.466
	PENJ		.481	2.438*	.030
	SI		-.014	-.084	.934
	Trust		-.125	-.594	.563
	Model: R^2=.71; F=5.410; df=6,13; P=.005				

Table 5.6. Regression analysis on Intention to Use comparing the conditions.

Testing the moderating factor hypotheses (Section 4.4) with a Chow test leads to the results presented in Table 5.7. It shows only Hypotheses 2a-c and 5b because only these could be tested: the other hypotheses concerned construct interrelations that were not significant in the regression analysis (Table 5.5).

Hypothesis	Factor*Variable	Dependent	F	Sig.
MFH2a	Gndr*PEOU	ITU	1.454	.236
MFH2b	Age*PEOU	ITU	2.498*	.024
MFH2c	EXP*PEOU	ITU	.808	.587
MFH5b	EDU*ANX	PEOU	1.969	.086

Table 5.7. Chow's test on moderating factor hypotheses

There was only one significant moderating factor: age on Perceived Ease of Use, indicating that the older the participants were, the more they found Ease of Use to be determining their Intention to Use.

5.3.3 Model test conclusions

Our results suggest a clear path: a robot with more social abilities has a higher score on Perceived Sociability, which contributes to a higher score on Social Presence. The latter contributes to a higher score on Perceived Enjoyment which contributes to a higher Intention to Use this system.

We may conclude that the sense of presence that people feel with a robot can be manipulated by changing its social abilities (which indeed makes people change their perception of social abilities accordingly). This sense of presence has a positive impact on the enjoyment that is felt and this is influencing its acceptance. In other words: the more a robot feels like a real person, the more fun it is - and the more fun it is, the more it is intended to be used.

Figure 5.4 shows a visualization of the results of the regression analysis. It shows that acceptance is merely dependent on Perceived Ease of Use and

Perceived Enjoyment. The users appreciate this system not so much for its usefulness, but for the enjoyment it provides and the easy way it can be controlled.

An important finding is the very small difference in R^2 value for the two conditions, despite the different scores. It shows that the social constructs provide a balanced model in this specific context. However, to draw decisive conclusions that can be generalized on the R^2 values, we have to await the results from experiments on different systems.

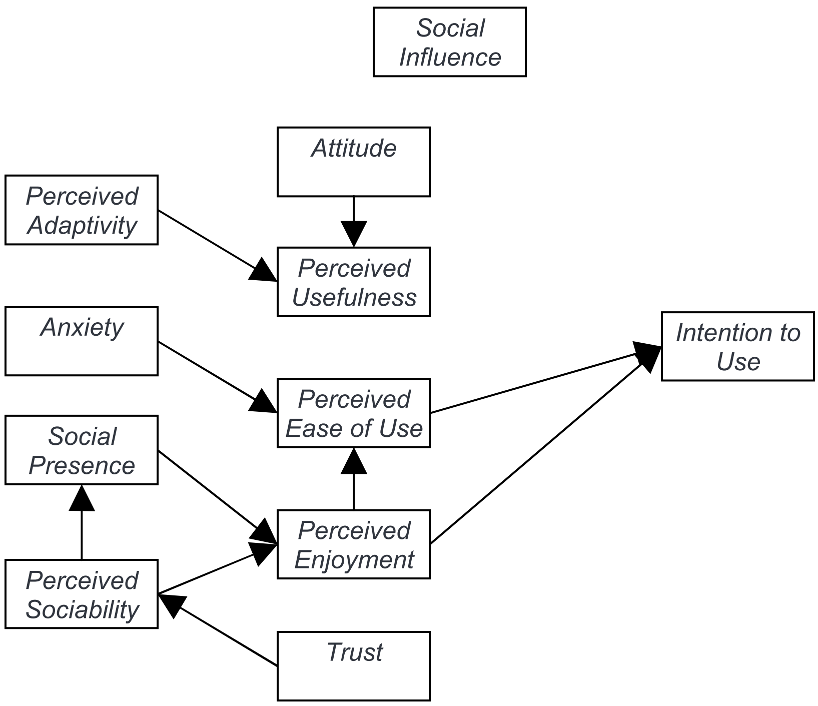


Fig. 5.4. Confirmed hypotheses concerning the model validation hypotheses

5.4 Observing conversational expressiveness

5.4.1 Introduction

In our first experiment as described in chapter 3, we found that conversational expressiveness differed between the two conditions. We used a simple method to register observed behavior and did not link the behavior analysis results directly to the scores on the different constructs.

With our new model it would be interesting to investigate these links, especially concerning the new construct of social presence. In human communication, conversational partners show a higher rate of expressivenss when a stronger social presence is perceived (Wagner and Smith 1991; Lee and Wagner 2002). The rate of expressivens demonstrates the amount of conversational engagement someone feels (Nakano and Nishida 2005). We call this conversational expressiveness: the amount and intensity of facial expressions and gestures when engaged in a conversation. We hypothesize that also for our user group, a higher score on the construct of Social Presence will correlate with a higher score on Conversational Expressiveness. As a higher score on Social Presence correlates with a higher score on acceptance (as indicated by the indicated Intention to Use the system), we suspect Conversational Expressiveness to correlate with Intention to Use.

5.4.2 Behavior analysis methodology

Although participants were observed during the experiment, we based our analysis on observations of the video's afterwards.

During the analysis non-verbal forms of conversational expressiveness were counted for each participant such as greeting the robot nodding or shaking the head, smiling, looking surprised (raising eyebrows) or irritated (frowning), and moving towards or away from the robot. This list of items considering conversational expressiveness was generated by listing classical feedback gestures (see Scherer 1987; Cerrato 2002; Axelrod and Hone 2005; Sidner and Lee 2005; Heylen et al. 2006)) without categorizing them to specific communicative functions (see Section 3.6.4 and the left column of Table 5.8).

To each counted item, the observers attributed two values: one for the strength (weight) of it and one for the certainty of the observer. Both could be one, two or three points. So if the observer would be sure of someone laughing very loud, this would score two times three points.

The observers watched the video's in which the camera was turned towards the participant, so the robot was not visible. They where not made aware of the different conditions of the robot. We had two observers for each video.

5.4.3 Analysis

To measure conversational expressiveness the occurrences of expressive behavior were counted and categorized for each participant. To account for inter-rater reliability we had to take into account that for every behavior, observants could or could not agree on seeing it and on their rating of certainty and weight. For this multidimensional type of rating often used measures as percentage of agreement and Cohen's kappa are not recommended (Banerjee et al. 1999). Moreover using solely correlation - either Pearson or Spearman - is usually not recommended, because neither coefficient takes into account the magnitude of the differences between raters. A correlation based method that compensates for this, is Lin's Concordance (Lin 1989; Lin 2000). We calculated the concordance score for our results (see Table 5.8), and found it to be 0.944 on average. Subsequently the scores were added.

Table 5.8 shows that there is no significant difference between the conditions.

	Mean	Standard Deviation	Lin's Concordance	t	Sig. (2-tailed)
Nodding head	2.28	3.004	0.940	-1.903	.065
Shaking head	1.13	2.078	0.958	.834	.410
Greeting	.63	1.531	0.970	-.102	.919
Lifting shoulders	1.53	2.736	0.983	-.515	.609
Suddenly moving away	.45	1.260	0.936	.497	.622
Suddenly approaching	.90	1.766	0.974	.354	.725
Smile	3.15	3.534	0.980	-1.646	.108
Laugh	3.33	4.492	0.980	-.877	.386
Raise eyebrows	.43	1.412	0.939	.111	.913
Frown	.15	.662	0.783	.000	1.000
Average			0.944		

Table 5.8 – Means, concordance and t scores on items of conversational expressiveness with t-scores comparing a more and less social condition

Table 5.9 shows that there is a correlation between Social Presence and Conversational Expressiveness (CE). There is no correlation however between Intention to Use and conversational expressiveness, which means we have no reason to assume that the score on conversational expressiveness is related to acceptance.

As in our previous behavior analysis study, we categorized the behavior types as positive and negative, and looked at the total number of times a type of behavior (positive/negative) occurred for the different conditions. We considered the behaviors *shaking head move away* and *frown* negative and all others positive. Table 5.9 shows that behaviors categorized as negative in fact did correlate with Intention to Use (of course in a negative direction).

		ITU	SP	CE
ITU	Correlation	1	.387*	.092
	Sig. (2-tailed)		.014	.574
SP	Correlation	.387*	1	.331*
	Sig. (2-tailed)	.014		.037
CE	Correlation	.092	.331*	1
	Sig. (2-tailed)	.574	.037	
Pos	Correlation	.209	.378*	.954**
	Sig. (2-tailed)	.196	.016	.000
Neg	Correlation	-.359*	-.103	.289
	Sig. (2-tailed)	.023	.527	.071

Table 5.9 - Pearson correlations for constructs and conversational expressiveness

Table 5.10 shows that there is a clear difference between the more social and less social condition both in total expressions (CE) and in the total amount of expressions that can be categorized as positive expressions.

	Means	t	Sig. (2-tailed)
Positive	14.475	3.058**	.004
Negative	1.525	-.502	.619
All CE	16.000	2.706*	.010

Table 5.10 - Means and t scores on categorized items of conversational expressiveness comparing a more and less social condition

5.4.4 Behavior analysis conclusions

Concerning our behavior observation, there is a clear pattern of more conversational expressiveness, a higher frequency of (positive) non-verbal behaviors of participants that were in conversation with the robot in a more social condition. This corresponds with a higher score on Social Presence, showing users experiencing a social entity are indeed responding to that.

This may say something about the effect of what we understand as Social Presence on users, but although Social Presence correlates with Intention to Use, Conversational Expressiveness (CE) only partly seems to be related to acceptance: an increasing amount of shakes, frowns and taking distance may indicate a lower acceptance rate.

5.5 Discussion

It would be interesting to see if the conclusions on the results of experiment 3 (Section 5.3.3) are specific for this type of robot and for elderly users or can be generalized. Future research could focus on different robots (and perhaps screen agents) and user groups, but also on the different ways this experience can be optimized.

Regarding the behavior observation we think that this can be an additional instrument for studies on robot acceptance. Further research could explore its possibilities and establish how it can be related to other data. A detailed discourse analysis for example, could provide clues that can be related to acceptance, although a different (non Wizard of Oz) setup would in that case be more appropriate.

An item for further research could be the question whether conversational expressions occurred in response to the same expressions by the robot (a smile in response to a smile, a frown in response to a frown). In that case we would be speaking of imitative behavior. This would be the occurrence of a well known phenomenon in psychology called the chameleon effect (Chartrand and Bargh 1999). It concerns imitative behavior between humans, which seems to occur naturally unless two people do not like each other.

The occurrence of this behavior could even very well be interpreted as a sign of acceptance (Kahn et al. 2006). But during behavior analysis the observers just counted the number of behaviors, without looking at the behavior of the robot that evoked it - the camera was always directed towards the participant. In future research this possibility of imitative behavior could be something to observe, also when comparing agents with different embodiments, since it could add interesting viewpoints to HRI theory on this aspect (Dautenhahn and Nehaniv 2002).

6. Exploring adaptiveness, adaptability and user control

Parts of this chapter have been published earlier in (Heerink et al. 2008c; Heerink et al. 2008d) and (Heerink et al. 2010b)

6.1 Introduction

As technological advances make it possible for systems to respond to users with more flexibility and autonomy, it becomes more common for these systems to adapt or be adapted. For some systems this concerns user-adaptation, possibly by learning from interaction or by detecting the specifics of a user (Benyon 1993; Cheverst et al. 2005). For context-aware systems it means gathering information from the environment to adapt themselves to the current situation (Scholtz et al. 2004; Schmidt 2005). These developments lead to adaptive applications in many different domains, including shopping recommenders that direct consumers to products that may be of interest to them (Alpert et al. 2003), mobile agents monitoring the user's surroundings in crisis situations (Streefkerk et al. 2006) and personalized tours (Fink and Kobsa 2002; Wubs and Huysmans 2006).

However, for aging adults, adaptive technology has its own requirements and perspectives. As we discussed in section 4.4.3, it is essential for our specific user group of older adults that assistive devices be either adaptive (self-adapting) or adaptable (can be adapted) because of the changing needs of the users (Pew and Hemel 2004). Growing older is a process during which physical and mental functions of our bodies gradually become less usable, due to which we need help, either in the form of humans or in the form of assistive devices. Older adults usually want neither humans nor devices to help them out when this help is not yet needed. They do not want a device to help them remember things as long as their memory still (more or less) functions and they do not want to be helped walking as long as they can still manage to walk by themselves. It is appreciated however, if these devices or people become helpful as soon as help is needed. This is partly to postpone the use of these devices because they could be stigmatizing (see for examples Forlizzi et al. 2004) and partly because the users want to keep their independence and remain using their physical and mental capabilities as long as possible (Jorge 2001; Ebersole et al. 2003).

This makes adaptivity a reoccurring requirement in projects concerning eldercare technology in general (Yu et al. 2003; Miller et al. 2004; Pew and Hemel 2004) and more specific in robot and screen agent technology (Kawamura et al. 2003; Maciuszek and Shahmehri 2003; Forlizzi et al. 2004). We therefore introduced the construct of Perceived Adaptivity in our model (Section 4.4.3). However, as a construct within an acceptance model, Perceived Adaptivity is completely unprecedented. In a similar way as we designed an experiment that could justify the ‘social constructs’ in the previous chapter, we therefore designed an experiment that could justify the addition of Perceived Adaptivity. This means we want to establish whether a more adaptive robot is accepted better than a less adaptive robot.

There is however another question that needs to be answered, concerning both the interpretation of the construct and the response to adaptive technology. This response is not always positive when it concerns systems that autonomously adapt to the user or the environment. Especially when these systems become more sophisticated, they perform actions that users never experienced from similar systems before (Höök et al. 2000), and this makes these systems to be perceived as unpredictable and unreliable (Höök 1998; Jameson 2003).

As Dautenhahn (2004) points out, there are two views in HRI on this that appear contradictory. On the one hand, there are indications that more autonomy would lead to more useful agents (Maes 1994) while on the other hand, there are indications that predictability and controllability should prevail (Shneiderman 1997). As we can generally state that adaptivity potentially makes the user feel no longer in control, the question is: should a system therefore be less adaptive? Should it rather be adaptable, or perhaps adaptive but with a form of user control? Should a system ask for confirmation before it autonomously adapts?

Several studies addressed these questions, finding that indeed the desire for user control limits the acceptance of autonomy (Gillies and Ballin 2004; Marble et al. 2004; Price et al. 2005), which means there is a delicate balance between automation/autonomous behavior and user control. We want to know how this intervenes with the interpretation and perception of adaptivity by our target group.

The statements that are used in our questionnaire (see Table 6.1) cover two concepts that are related, but nonetheless refer to different underlying processes: *adaptability* and *adaptiveness.*

PAD	*14. I think the robot can be adaptive to what I need*
	15. I think the robot will only do what I need at that particular moment
	16. I think the robot will help me when I consider it to be necessary

Table 6.1. Statements for the construct of Perceived Adaptivity (PAD)

The items can be interpreted as either the user being able to adapt a device or system to his or her demands or needs, or the system adapting autonomously. In our case both adaptive changes – the first by the user and the second by the system itself – are supposed to be related to the gradually growing need for assistance by an older adult. Furthermore, the questionnaire items do not suggest either presence or absence of user control: adaptiveness could for example be performed either with or without asking this user for approval. In the latter case we speak of user control, which is lacking when a system adapts without asking for approval.

Thus we have the following options to further specify the PAD-construct:

- adaptable: the user adapts the robot to his or her changing needs;
- adaptive with user control: the robots adapts to observed changing needs of the user after the user has agreed to this;
- adaptive without user control: the robot adapts to observed changing needs of the user without seeking agreement of the user.

Summarizing we state that in this chapter we will address two issues: confirming the relevance of the construct of Perceived Adaptivity by comparing a more adaptive version of a robot to a less adaptive one and solving the ambiguity of the interpretations of Perceived Adaptivity.

This means we want answers to the following questions:

- Does a robot that is more adaptive/adaptable have a higher score on Intention to Use than a less adaptive robot?
- Is there a difference considering the scores for Perceived Adaptivity and Intention to Use between an adaptive and an adaptable robot?
- When adaptive/adaptable, is there a difference in impact on the scores for Perceived Adaptivity and Intention to Use if there is more user control?
- Does a request for confirmation lead to a higher sense of control by the user?
- Is a more adaptive/adaptable robot found to be more useful than a less adaptive/adaptable one?

We will address these questions by rephrasing them in the form of hypotheses which we will test by carrying out an experiment using a video of an elderly user with an assistive social robot. It has four conditions: a neutral one, an adaptable one, an adaptive one with user control and an adaptive one without user control.

6.2 Focus within the model

Our focus for this chapter is on Perceived Adaptivity. Furthermore, since a more adaptive/adaptable system is expected to make this system more useful, we also focus on the relationship between Perceived Adaptivity and Perceived Usefulness, and on how this influences the Intention to Use. This means we are focusing as visualized in Figure 6.1.

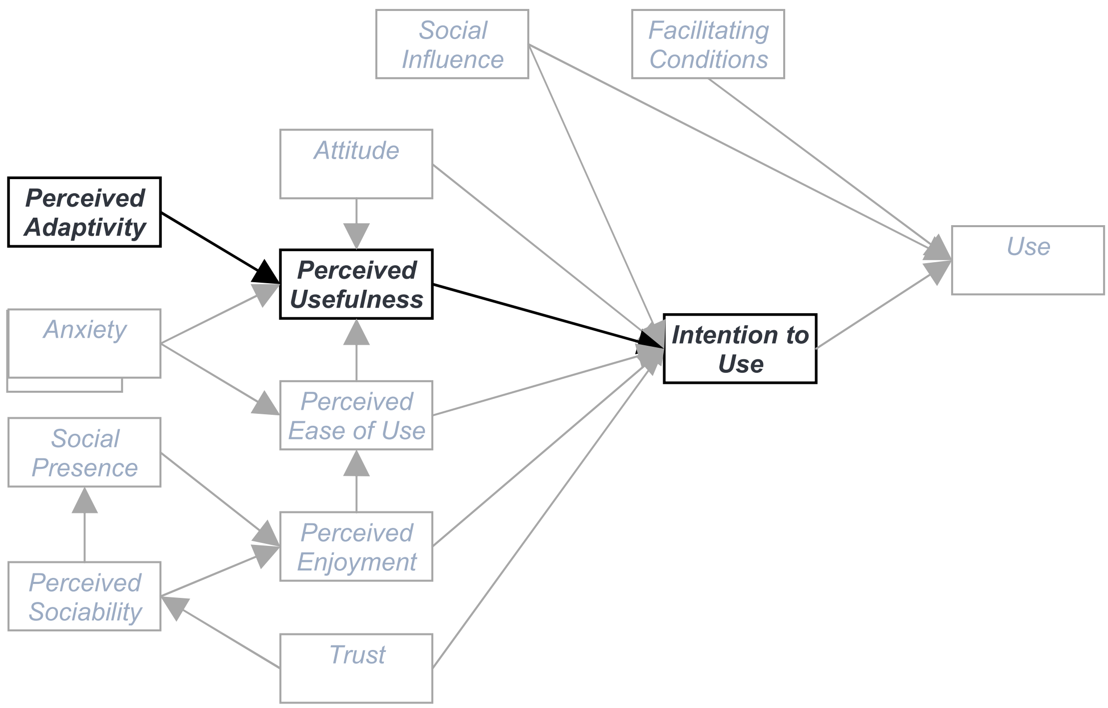


Figure 6.1. Focus of the experiment

We formed hypotheses to be tested, based on the questions we described in section 6.1 and on the overall assumptions concerning the relevant construct interrelations of our model. The first two hypotheses belong to the latter category, demanding a confirmation of assumed construct interrelations:

Exp4-H1 Intention to Use is determined by Perceived Usefulness.

Exp4-H2 Perceived Usefulness is determined by Perceived Adaptivity.

The third hypothesis assumes that our manipulations will be reflected in the score for Perceived Adaptivity:

Exp4-H3 The score on Perceived Adaptivity will be higher for the adaptable and both adaptive conditions compared to the condition that is neither adaptable nor adaptive.

Also, we expect the user control condition to have a higher score on related statements:

Exp4-H4 Participants will indicate to sense more user control in a condition where this is implemented.

Furthermore, we are interested in differences between the adaptable and adaptive conditions. As it is generally found that adaptiveness increases accessibility of assistive technology for elderly users (Jorge 2001; Miller et al. 2004; Pew and Hemel 2004), we expect that the Perceived Adaptivity of the adaptive versions leads to a higher score on Perceived Usefulness. This would subsequently affect Intention to Use.

Exp4-H5 There will be a higher score on (a) Perceived Adaptivity (b) Perceived Usefulness and (c) Intention to Use for both adaptive conditions as compared to the non adaptive conditions.

Moreover, we expect a difference between the two adaptive conditions with and without user control. As we stated in the previous section, the desire for user control limits the acceptance of autonomy: in general users seem to feel less anxiety for a robot if they experience more user control, we expect the user control condition to be preferred. This would be reflected in higher scores on Anxiety (this means less anxiety since the scores are reversed) and through Perceived Usefulness on Intention to Use

Exp4-H6 There will be a higher score on (a) Anxiety (b) Perceived Usefulness and (c) Intention to Use for a condition with user control as compared to a condition without user control.

6.3 Method

We wanted to set up an acceptance measuring experiment in which we could have an assistive social robot in four conditions: a neutral one that is neither adaptable nor adaptive, an adaptable one, an adaptive one with user control and an adaptive one without user control. To effectively compare these four conditions, we needed at least 20 participants for each condition, which meant we needed a group of at least 80 participants.

To meet this challenge, we decided to use a video of a robot interacting with an elderly actor instead of a real live HRI trial to create the four conditions. Using video's in HRI trials is found to be a method that leads to results that are comparable to live trials (Woods et al. 2006a; Woods et al. 2006b).

6.3.1 System

We were able to use video material made for the Robocare project by the Institute for Cognitive Science and Technology of the Italian National Research Council for research by Cesta et al. (Cesta et al. 2007; Cesta et al. 2007).

The RoboCare project concerns a service type robot. It is cylinder shaped and mobile (wheels), and it is connected to a system that features sensors and cameras. It has the possibility to produce pre-programmed speech. There is a version with a screen on which a female face is displayed to embody the conversation. The robot serves both as an interface to 'smart home' technology and as an autonomous actor, retrieving information from it's intelligent environment and acting upon this. The RoboCare project is not so much focused on developing a robot as to an environment, an intelligent home of which a robot is an integrated part (Bahadori et al. 2003). Published research related to

RoboCare is focused on technical matters or design issues – for example comparing responses to a robot with a screen, a face or just a voice.

We made four video's of the robot (see screenshots in Figure 6.2), representing the four conditions. In all these videos, the robot had the same three functionalities which were already presented in the original video as developed by the Robocare researchers (the original video is available online at http://robocare.istc.cnr.it/videos/rbc-sample-1.avi):

1. monitoring the user and alarming if necessary;
2. helping to remember to take the right medication at the right time;
3. functioning as a fitness advisor (announcing that it is time for some exercise if the user has been seated too long).

In the first, neutral condition, the robot simply had al these functionalities: the user could not turn them on or off and the system did not modify them by itself. In the second, adaptable condition, the second functionality was shown to be turned on by the user. This function was most suitable to be the adaptable/adaptive feature: as we reported earlier, the reminder function could be something that made participants reject the use of a robot as long as they felt their memory was still good enough.

Figure 6.2. Stills from the video's

In conditions three and four, both adaptive, the second functionality (medication reminding) was turned on by the system itself. In the third condition, there would be user control: the system would suggest the functionality to be turned on and would await the user's approval before doing this. In the fourth condition there would be no user control: the system would simply announce the functionality to be necessary and turn it on.

So these were the four conditions, represented by four videos:
1. Not adaptive, not adaptable.
2. Adaptable, not adaptive.
3. Adaptive, not adaptable, with user control.
4. Adaptive, not adaptable, without user control.

6.3.2 Participants

We found 100 older adults willing to take part in the experiments who were living in apartments close to or within eldercare institutions in the cities of Almere and Amsterdam. Due to incomplete questionnaires and other procedural irregularities, we had to omit 12 participants from the results. So our results feature 88 participants, from which 28 were male and 60 were female (which is in accordance with the demographic overrepresentation of women in this age group for this generation). Table 6.3 lists their age, education and computer experience.

	Minimum	Maximum	Mean	Std. Deviation
Age	58	99	77.73	8.727
Experience	1.00	5.00	1.966	1.266
Education	2	10	4.78	2.120

Table 6.3. Descriptive statistics for age, computer experience and education

6.3.3 Procedure

There were three researchers who had all four videos on a laptop. They visited the participants, explained the set up of the experiment and showed one of the videos at each visit. So every participant just saw one video and the link of a participant to a particular video was randomly made. After this, the participant would be asked to fill out the questionnaire. If any help reading the form was needed, it would be given, but to avoid influencing the participants, the researchers gave no explanation.

6.3.4 Questionnaire adaptation

Although we are focusing on just a part of our model, we used the complete questionnaire of the new model (as presented in Table 4.6). This would enable us to compare the results to those of our previous experiments and even add the cases of all experiments for which the same questionnaire was used (this will be carried out in Chapter 8).

We added a control question to enable us to check whether the different versions would reflect the way the users perceived the robot. We made this a multiple choice question with four answers – answer a. corresponded with the first version, answer b. with the second one and so on (see Table 6.4).

We also introduced a user control statement, saying that the user in the video had control over the robot. As with the regular questionnaire items, this could be replied to on a five point scale.

Category	**Statement/question**
Manipulation check	What happened in the last scene? a. The robot reminded the women that it was time to take her medication. b. The robot reminded the women that it was time to take her medication after she turned the option 'medication reminder' on. c. The robot told the women that she had taken her medication in time and asked if he should remind her next time. d. The robot told the women that she had taken her medication in time and that he would remind her next time.
User control	The woman in the video controls what the robot does and does not do.

Table 6.4. Added statements for user control and manipulation check question

6.4 Results

The 88 questionnaire forms that turned out to be usable had the following numbers of participants divided over the four video's (Table 6.4).

Condition	***Description***	***N***
1	*Not adaptive, not adaptable*	*22*
2	*Adaptable*	*21*
3	*Adaptive user controlled*	*23*
4	*Adaptive not user controlled*	*22*

Table 6.4. Robot conditions and number of participants

Table 6.5 shows the descriptive statistics for the combined scores on the conditions.

	Minimum	**Maximum**	**Mean**	**Std. Dev.**
User Control	1	5	3.32	.977
Anxiety	1.25	5.00	3.671	.734
Attitude	1.33	5.00	3.167	.922
Facilitating Conditions	1.50	5.00	3.477	.7386
Intention to Use	1.00	5.00	3.402	1.052
Perceived Adaptivity	1.33	5.00	3.492	.645
Perceived Enjoyment	1.50	5.00	2.955	.860
Perceived Ease of Use	1.60	5.00	3.559	.714
Perceived Sociability	1.00	4.25	2.696	.645
Perceived Usefulness	1.00	5.00	3.633	.893
Social Influence	2.00	5.00	3.205	.730
Social Presence	1.00	4.40	2.830	.533
Trust	1.50	5.00	3.602	.898

Table 6.5. Descriptive statistics for combined conditions

Calculating Cronbach's Alpha (Table 6.6) we found that all of the constructs were reliable, except for Facilitating Conditions, which also was not reliable if one of the questions were omited. Since we hypothesized (see section 4.4.1 and 4.5) this construct only to relate to actual use, which was not measured in this experiment, this was not a relevant construct for our hypotheses. We therefore omitted it from the results of this experiment. For the other constructs we had to omit statements 21 (PENJ5), 22 (PEOU1), 37 (SI3) and 45 (Trust3). These were also omitted in our previous experiment.

Construct	**Alpha**	**Construct**	**Alpha**
Anxiety	*.701*	*Perceived Ease of Use*	*.842*
Attitude	*.763*	*Perceived Sociability*	*.717*
Facilitating Conditions	*.426*	*Perceived Usefulness*	*.825*
Intention to Use	*.854*	*Social Influence*	*.701*
Perceived Adaptivity	*.792*	*Social Presence*	*.735*
Perceived Enjoyment	*.756*	*Trust*	*.758*

Table 6.6. Cronbach's Alpha

To establish the strength of the association between the video versions and the manipulation check (MC) question (Table 6.4) we generated the cross tabulation which is presented in Table 6.7. The significance of this relation can be established by calculating Cramers V. This is a chi-square-based measure of nominal association which gives a normalized value between 0 and 1 (Cramér 1999). In our case, the value for Cramers V is .714, which is significant at the 0.001 level. This means there is a strong association between the manipulation check question and the video versions: participants generally perceived the amount of adaptability, adaptivity and user control that was consistent with the video version they had seen.

		MC question				
		1	**2**	**3**	**4**	**Total**
Video	1	21	1	0	0	22
	2	3	16	1	1	21
	3	6	2	12	3	23
	4	1	0	2	19	22
Total		31	19	15	23	88

Table 6.7. Cross tabulation MC question and Video

To establish the effect of our manipulations further, we compared the results of the participants that saw the first video (neither adaptable nor adaptive) to the scores related to the other three video's that concerned either adaptable or adaptive conditions. Table 6.8 shows that Perceived Adaptivity indeed scored much higher for the video's that featured an adaptable or adaptive robot (M=3.661, SD=.550 versus M=2.984, SD=654). Also Perceived Usefulness scored higher for these video's (M=3.742, SD=.882 versus M=3.303, SD=860), but this did not lead to a higher score on Intention to Use. This means Hypothesis 5 is partly confirmed.

variable	t	Sig.	variable	t	Sig.
Anxiety	.417	.677	Perceived Sociability	1.763	.082
Attitude	1.432	.156	Perceived Usefulness	2.035*	.045
Intention to Use	1.133	.260	Social Influence	2.632*	.011
Perceived Adaptivity	4.762**	.000	Social Presence	.835	.407
Perceived Enjoyment	.606	.547	Trust	1.591	.115
Perceived Ease of Use	1.565	.121	User Control	2.988*	.004

Table 6.8. T-test comparing video 1 related scores (neither adaptable nor adaptive) to the scores related to the other video's (adaptable/adaptive)

We subsequently compared the non adaptive conditions to the adaptive conditions. The first category consisted of the scores for video 1 and video 2 and the second of video 3 and 4. Table 6.9 shows the t-test scores to compare these two condition sets. Perceived adaptivity was indeed higher in the adaptive condition set (M=3.674, SD=.566) compared with the non-adaptive condition set (M=3.302, SD=.673). This indicates our manipulation was successful.

Also, participants who viewed the adaptive robot video were found to have a higher score on Intention to Use and a more positive attitude toward the robot. They percieved the robot more as enjoyable and more useful. Furthermore, participants in the adaptive condition reported to feel less anxiety toward the robot and found other people's opinions about using the robot (social influence) more important.

variable	t	Sig.	variable	t	Sig.
Anxiety	2.334*	.022	Perceived Sociability	.306	.761
Attitude	3.023*	.003	Perceived Usefulness	3.523*	.001
Intention to Use	3.485*	.001	Social Influence	2.178*	.032
Perceived Adaptivity	2.807*	.006	Social Presence	.350	.727
Perceived Enjoyment	2.298*	.024	Trust	1.655	.102
Perceived Ease of Use	1.150	.254	User Control	2.170*	.033

Table 6.9. T-test comparing adaptive to non adaptive

To compare the four conditions represented by the four video versions, we used a one way ANOVA (see box plots Figure 6.3 and 6.4), accompanied by a post hoc Games Howell comparison analysis as shown in Table 6.10. A Games Howell comparison (Games and Howell 1976) is a usual instrument in cases where multiple groups have to be compared pair wise (Zwick 1986). As with t-test results, a positive value means a higher score for the first of the two compared groups and a negative score means a higher score for the second group. It does not require equal variances.

Remarkable in these results is the distinguishing score of the third condition in relation to the other ones. It scores higher than the first condition on User Control, Perceived Adaptiveness and Perceived Usefulness, and it scores higher than the second condition on Anxiety, Intention to Use, Perceived Enjoyment and - again - Perceived Usefulness. An adaptive robot that asks for confirmation (condition 3) is thus clearly perceived as more useful than a robot that is not adaptive.

	2 to 1	Sig.	3 to 1	Sig.	4 to 1	Sig.	2 to 3	Sig.	2 to 4	Sig	3 to 4	Sig.
UC	.377	.430	.960*	.003	.273	.700	-.584	.234	.104	.299	.688	.130
ANX	-.256	.655	.397	.278	.057	.989	-.654*	.049	-.313	.445	.340	.359
ITU	-.325	.770	.833	.055	.318	.709	-1.157*	.002	-.643	.106	.514	.176
PAD	.650*	.005	.856*	.000	.515*	.044	-.206	.545	.135	.860	.341	.176
PENJ	-.268	.615	.582	.078	-.034	.999	-.849*	.011	.234	.254	.616	.114
PEOU	.236	.645	.393	.186	.182	.863	-.157	.850	.054	.996	.211	.791
PU	.014	1.000	.885*	.009	.379	.393	-.871*	.009	-.364	.411	.507	.176

Table 6.10. Games-Howell comparison between the four conditions

Moreover, the adaptive version without user control (4) only scores significantly higher on Perceived Adaptiveness when compared to the first condition – it does not score higher on any other construct, compared to any other condition. Also, there is no significant difference between the two adaptive conditions (3 and 4). Apparently, only the combination of adaptiveness and user control can make a clear difference in user perception.

The plots shown in Figure 6.3 to 6.6 confirm the outstanding scores for the adaptive condition with user control, showing the third condition with the highest score on Intention to Use (Figure 6.3), Perceived Usefulness (Figure 6.5) and on User Control (Figure 6.6). Furthermore, the plot in Figure 6.4 shows what is also very clear in Table 6.10: both adaptability and adaptiveness lead to a higher score on Perceived Adaptiveness.

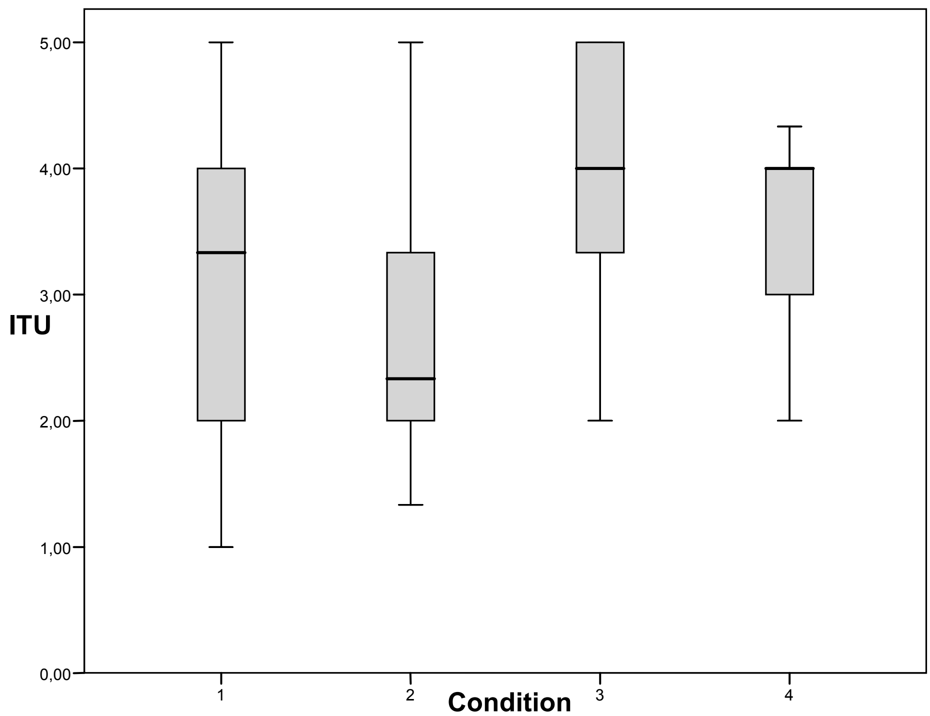


Figure 6.3 Box plot for Intention to Use

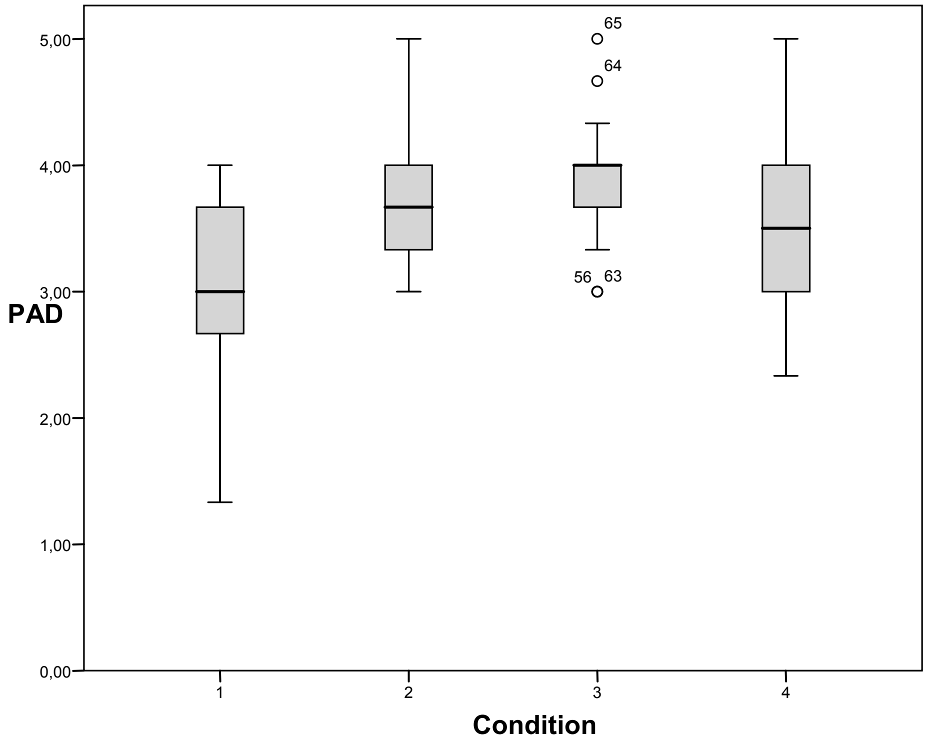


Figure 6.4 Box plot for Perceived Adaptivity

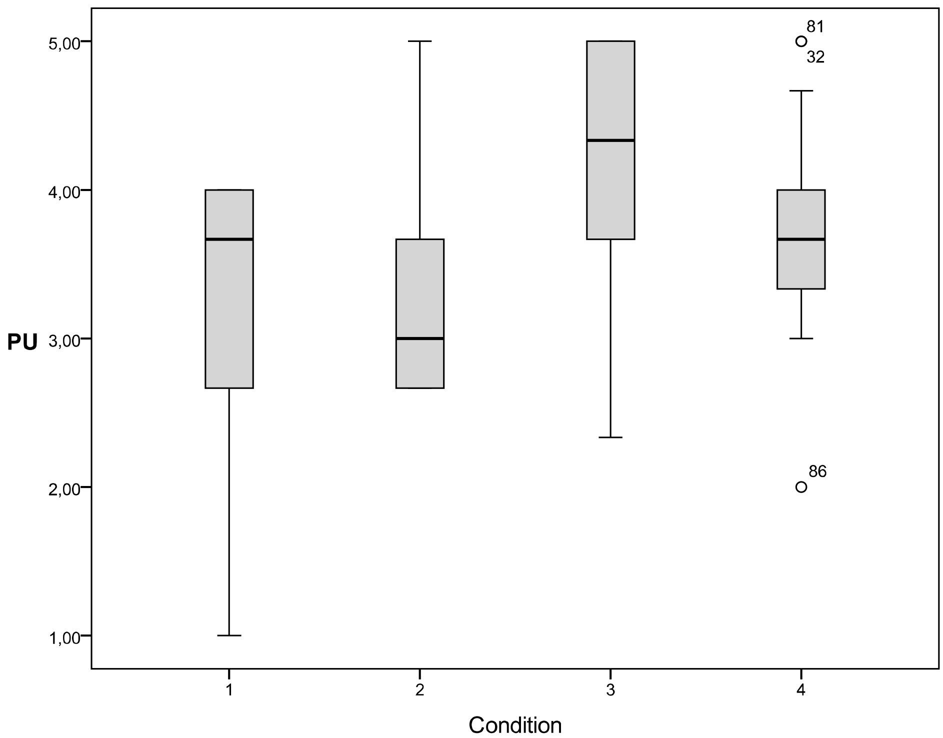


Figure 6.5 Box plot for Perceived Usefulness

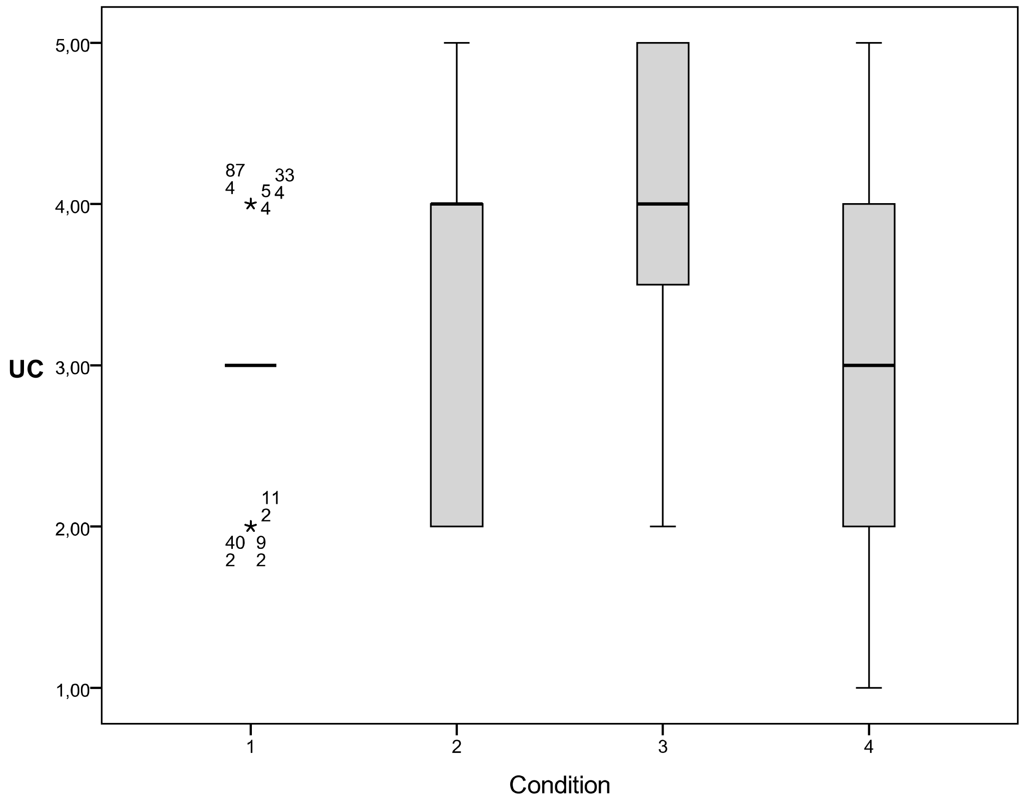


Figure 6.6. Box plot for the user control question

Table 6.11 shows correlation scores concerning the constructs that are relevant here plus the score on the user control question (a complete list can be found in appendix E). Perceived Adaptivity correlates with Perceived Usefulness. As can be expected, the score on user control (UC) correlates with the construct of Anxiety: the more user control is perceived, the less anxiety is experienced.

		UC	ANX	ITU	PAD	PENJ	PEOU	PU
ANX	Corr	.372**	1	.188	-.009	.286**	.436**	.151
	Sig.	.000		.079	.931	.007	.000	.161
ITU	Corr	.008	.188	1	.373**	.531**	.403**	.718**
	Sig.	.938	.079		.000	.000	.000	.000
PAD	Corr	.162	-.009	.373**	1	.196	.280**	.338**
	Sig.	.132	.931	.000		.067	.008	.001
PENJ	Corr	.175	.286**	.531**	.196	1	.454**	.525**
	Sig.	.103	.007	.000	.067		.000	.000
PEOU	Corr	.223*	.436**	.403**	.280**	.454**	1	.428**
	Sig.	.037	.000	.000	.008	.000		.000
PU	Corr	.193	.151	.718**	.338**	.525**	.428**	1
	Sig.	.072	.161	.000	.001	.000	.000	

Table 6.11. Correlations between major items

We also tested the model validation hypotheses (Section 4.5) – again without H1 which concerns usage – with a regression analysis on the results for the combined conditions. Table 6.12 shows that Perceived Usefulness and Attitude are the main determining influences on Intention to Use. This means the strong utilitarian aspect of this robot type (it does not entertain like the iCat in the previous experiment) is reflected in the scores. Furthermore, it shows that Social Presence is not a determining influence on Perceived Enjoyment. This could have to do with the relatively low score of Social Presence (2.830 versus 3.600 in the previous experiment), which is not remarkable if we consider the fact that participants did not experience a real robot. Still Perceived Sociability did influence Perceived Enjoyment, indicating that the (lower scoring) abilities were still of influence on the way the robot was perceived, although this had no impact on Intention to Use.

Table 6.13 shows an analysis of the performance of the model for the four different conditions. It shows a relatively high R^2 score for the adaptable condition, which could be related to the strongly determining influences of Perceived Usefulness and Attitude. This can be interpreted as the participants having a positive attitude towards the robot, finding the adaptability very useful.

Hypothesis	Independent	Dependent	Beta	t	Sig
H2 (a)	PU	ITU	.330	3.642**	.000
(b)	PEOU		.047	.624	.534
(c)	ATT		.496	5.442**	.000
(d)	PENJ		.031	.375	.709
(e)	SI		-.015	-.192	.848
(f)	Trust		.086	1.096	.276
	Model: R^2=.68; F=29.341; df=6,81; P=.000				
H3 (a)	ANX	PU	.047	.538	.592
(b)	ATT		.552	6.637**	.000
(c)	PAD		.185	2.251*	.027
(d)	PEOU		.365	3.264**	.002
	Model: R^2=.50; F=20.757; df=4,83; P=.000				
H4 (a)	ANX	PEOU	.334	3.543**	.001
(b)	PENJ		.359	3.809**	.000
	Model: R^2=.31; F=18.931; df=2,85; P=.000				
H5 (a)	PS	PENJ	.419	3.559**	.001
(b)	SP		.169	1.433	.156
	Model: R^2=.30; F=17.626; df=2,85; P=.000				
H6	Trust	PS	.571	6.445**	.000
	Model: R^2=.33; F=41.533; df=1,86; P=.000				
H7	PS	SP	.632	7.565**	.000
	Model: R^2=.40; F=57.231; df=1,86; P=.000				

Table 6.12. Regression analysis on model validation hypotheses

Figure 6.7 shows a model diagram in which the confirmed construct interrelations are visualized.

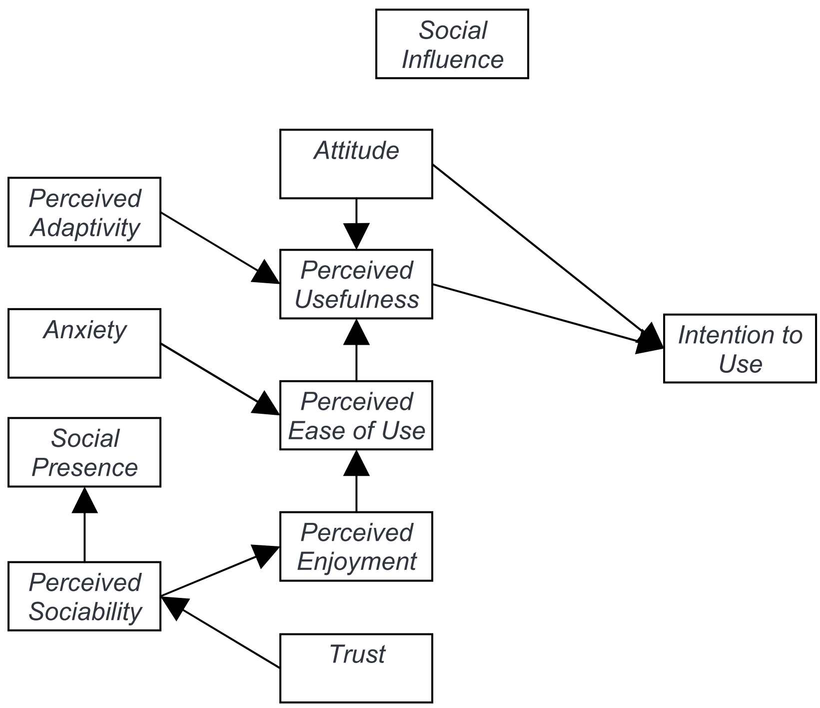


Figure 6.7. Confirmed relations

Condition	Independent	Dependent	Beta	T	Sig
Neutral	PU	ITU	.486	1.875	.080
	PEOU		-.033	-.153	.881
	ATT		.215	2.204*	.030
	PENJ		.098	.554	.588
	SI		.011	.064	.950
	Trust		.086	.413	.686
	Model: R^2=.69; F=5.676; df=6,15; P=.003				
Adaptable	PU	ITU	,598	3,015*	,009
	PEOU		-,116	-,604	,556
	ATT		,695	3,950**	,001
	PENJ		,137	1,213	,245
	SI		-,345	-1,572	,138
	Trust		,107	,640	,533
	Model: R^2=.88; F=16.445; df=6,14; P=.000				
Adaptive with user control	PU	ITU	,180	,853	,406
	PEOU		,235	1,335	,201
	ATT		,569	2,286*	,036
	PENJ		-,019	-,083	,935
	SI		,112	,512	,615
	Trust		-,107	-,574	,574
	Model: R^2=.69; F=5.974; df=6,16; P=.002				
Adaptive without user control	PU	ITU	,108	,547	,592
	PEOU		-,200	-,964	,350
	ATT		,642	2,854*	,012
	PENJ		-,142	-,585	,567
	SI		-,069	-,347	,733
	Trust		,450	1,971	,067
	Model: R^2=.62; F=4.129; df=6,15; P=.012				

Table 6.13. Regression analysis on ITU for the different conditions

Testing the moderating factor hypotheses (Section 4.4) with a Chow test leads to the results presented in Table 6.14. It shows only Hypotheses 1a, 1b and 5b: these were the only ones that could be tested, since the other hypotheses concerned construct interrelations that were not significant in the regression analysis (Table 6.12).

Hypothesis	Variable*Factor	Dependent	F	Sig.
MFH1a	Gndr*PU	ITU	1.454	.119
MFH1b	Age*PU	ITU	.913	.571
MFH5b	EDU*ANX	PEOU	.210	.988

Table 6.14. Chow's test on moderating factor hypotheses.

The table shows that we found no significantly moderating factors.

6.5 Conclusions

First of all it is remarkable that the adaptable condition (condition/video 2) is not accepted better, but both adaptive conditions (3 and 4) scored higher. Most clear however, are the outstanding results for the adaptive condition with user control (condition/video 3). We thus conclude that users prefer a system that adapts itself, requiring limited or no knowledge on operating it, but with still leaving the user in control.

Scores on Anxiety and Perceived Enjoyment also differ, especially between the second and the third condition, in favor of the adaptive one. This finding, combined with the results on the user control question, indicates that a robot that adapts itself after being permitted to do so, is more fun and less worrying. It is surprising however, that this does not lead to significant differences on the score for Perceived Ease of Use.

Looking back on our focus hypotheses (established in Section 6.2), we can confirm the first two construct interrelations: regression analysis shows Intention to Use is determined by Perceived Usefulness and the latter is determined by Perceived Adaptivity.

The third focus hypothesis can be confirmed. Not only did a t-test show clear differences between the first and the other conditions (Table 6.8), also the comparison scores (Table 6.10) show significant higher scores on Perceived Adaptivity for conditions 2, 3 and 4. Remarkably, scores for Intention to Use are the highest for the third version, but in our comparison analysis only the difference between the second and the third version is significant.

The fourth hypothesis cannot be fully confirmed: participants did indicate to sense more user control in a condition where this is implemented, but this is only true for the third condition (with user control) when compared to the first condition (see Table 6.10). Between the third and both second and fourth condition there is no significant difference although the score on user control was clearly the highest for the third condition (Figure 6.6). Our analysis did however show a strong correlation between user control and Anxiety (Table 6.11), indicating that an increased sense of control leads to anxiety decrease.

Furthermore, we can fully accept the fifth hypothesis: as the t-test results in Table 6.9 show, the scores Perceived Adaptivity, Perceived Usefulness and Intention to Use are clearly higher for the two adaptive conditions.

The last hypothesis has to be rejected, since (as Table 6.10 shows) the adaptive condition with user control (video 3) did not score higher on any construct than the adaptive version without user control (video 4).

This means we have found answers to the questions posed in section 6.1:

- We can state that in our experiment a robot that was more adaptive turned out to be more acceptable (i.e. leads to a higher score on Intention to Use) than a less adaptive robot, even if the latter concerned an adaptable robot. However, we could not establish an adaptable robot to be more acceptable than a non adaptable robot.
- When adaptive, a request for approval before adapting (suggesting more user control) leads to a higher score on acceptance.
- A request for approval by an adaptive robot (condition 3) does not directly lead to a higher sense of control by the user when compared to an adaptive robot that did not ask for approval (condition 4). However, the adaptive robot asking for approval (condition 3) scored significantly higher on user control than the non adaptive, non adaptable robot (condition 1).

We find that these results justify the addition of Perceived Adaptivity to our model. However, they also show that – indeed – there is a subtle balance between autonomous adaptivity and the desire for user control as we stated in the introduction of this chapter. Further research using similar measuring instruments could establish where this balance differs for different systems, user groups and perhaps stages in aging.

7. Usage experiments

Parts of this chapter have been published earlier in (Heerink et al. 2008; Heerink et al. 2008a; Heerink et al. 2008b) and (Heerink et al. 2010b)

7.1 Introduction

In chapter 4 we defined our hypothetical model and in chapters 5 and 6 we justified the addition of constructs that were unprecedented in technology acceptance modeling. In this chapter we focus on the validation of the model, establishing a relation between the Intention to Use as measured by our questionnaire and actual usage over a longer period (seven to ten days). For this purpose, we set up two experiments with different systems: a robot (again the iCat) and a screen agent (Steffie). For both studies we used the entire instrument, to be able to compare the results to our previous experiments. We will describe each experiment and its results separately and then evaluate the combined results.

7.2 Rationale and hypotheses

Our hypothetical model as presented in Chapter 4 consists of a set of constructs that supposedly either directly or indirectly determine the users' Intention to Use the robot. As usual in acceptance modeling, this means we assume this intention is a reliable indication of actual use (Lee et al. 2003). To establish whether this assumption can be made for this specific technology and user group, we have to be able to compare data on users' intention with data on their actual use of the technology. This means we have to set up an experiment that enables us to gather data on both Intention to Use and actual use. To do this, our first step can be to follow the same procedure as our previous experiments in which an initial introduction of a few minutes was followed by a questionnaire. After this, the system has to be available to the participants for a longer period of time. This way the participants can actually choose to use the technology during this period. Of course, it has to be registered how often and how long the participants are doing this.

As both the model overview and the constructs overview (see Figure 4.1. and Table 4.1.) show, the model incorporates the following hypothetical influences:

H1 Use is determined by (a) Intention to Use and influenced by (b) Social Influence and (c) Facilitating Conditions.
H2 Intention to Use is determined by (a) Perceived Usefulness, (b) Perceived Ease of Use, (c) Attitude, (d) Perceived Enjoyment, (e) Social Influence and (f) Trust.
H3 Perceived Usefulness is influenced by (a) Perceived Ease of Use (b) Attitude, (c) Perceived Adaptivity and (d) Anxiety
H4 Perceived Ease of Use is influenced by (a) Anxiety and (b) Perceived Enjoyment
H5 Perceived Enjoyment is influenced by (a) Social Presence and (b) Perceived Sociability
H6 Perceived Sociability is influenced by Trust
H7 Social Presence is influenced by Perceived Sociability

To test these hypotheses, we set up two experiments with different systems, used in different settings, each featuring 30 participants. These different settings and systems would give us some indications on the possibility to generalize the model and the predictability of usage by Intention to Use in particular. Both experiments were set up so that we could collect data on first impressions (as we did in the previous experiments) and on actual use of the systems. In order to keep a consistent numbering, we will refer to these experiments as Experiment 5 and Experiment 6.

The first experiment we will discuss, Experiment 5, featured again the iCat robot and was carried out in an eldercare institution where the robot could be used by anyone – so this could be considered 'actual use in a public setting'. The second experiment that will be described in this chapter, Experiment 6, featured a screen agent and was carried out at the homes of older adults that were living independently at their homes – a setting that could be described as 'actual use in a private setting'.

7.3 Experiment 5: actual use in a public setting

For Experiment 5 we used a setup in which the robot was connected to a touch screen as is shown in Figure 7.1. It could be used for information and for fun: the participants could ask for weather forecast, a television program overview or a joke by pressing the appropriate choices from a menu on the screen. The information was then given with pre-recorded speech by the iCat, for which we used a female voice. This speech was generated in Dutch with a text to speech engine (Loquendo[2]).

[2] We used the 2007 version of Loquendo (http://www.loquendo.com/en/technology/TTS.htm)

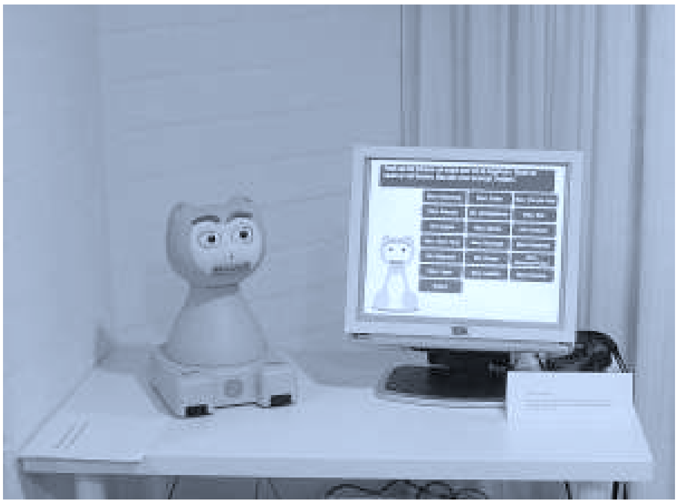


Fig. 7.1. Setup iCat with touch screen for Experiment 5

This experiment included a seven day period during which the system was available for use to anyone passing by. During this period, the system made video recordings as soon as it was used through the camera in iCat's nose: recording was launched by motion detection.

```
----------------------------------------------
De datum is: maandag 11 juni 2007
----------------------------------------------
iCat is gestart door deelnemer:
Mevr. Gerritsen
De starttijd is: 10:29:05.
De eind tijd is: 10:32:05.
iCat is gestopt door de gebruiker.
----------------------------------------------
----------------------------------------------
De datum is: maandag 11 juni 2007
----------------------------------------------
iCat is gestart door deelnemer:
Iemand Anders
De starttijd is: 10:33:23.
De eindtijd is: 10:34:14.
iCat is gestopt door de gebruiker.
----------------------------------------------
```

Fig. 7.2 iCat usage log example

Furthermore, the application kept a log that stated for each session the user name, the start and end times of individual user sessions. Figure 7.2 shows a fragment from this log. The end time was either the time a user actively ended his session or the time the system automatically ended the session if it was not used for 90 seconds.

7.3.1 Subjects

The experiment took place in an eldercare institution in Almere (the Netherlands). There were 30 participants (Table 7.1 shows their descriptive statistics), recruited both by eldercare personnel and by students. Their age ranged from 65 to 94, while 22 of them were female and 8 were male. 17 of them lived inside the eldercare institutions, 13 lived independently in apartments next to the institutions. As before, we asked both staff and students to approach those whose mental condition was such that a questionnaire could be coped with. Otherwise there was no selection on mental or physical health features.

	Minimum	Maximum	Mean	Std. Deviation
Age	65	89	79.37	7.19
Experience	1.00	5.00	2.13	1.63
Education	2	10	3.80	2.22

Table 7.1. Descriptive statistics for age, education and computer experience Experiment 5

7.3.2 Procedure

Participants were brought into a room were they were instructed to simply play with the robot for about three minutes, choosing any task they would like. Subsequently they were brought to another room where they were instructed to fill out the questionnaire. They could ask for help if they were unable to read the statements.

After these sessions were completed, we left the robot for public use in a tea room, for a seven days period. During this time, if the robot was not being used, the screen showed buttons with the names of the test session participants and one extra button saying “I’m not listed”. Passers by were informed by a note that anyone could use the robot and that they could start a session by pressing the button with their name on it or the “I’m not listed” button if their name was not on the screen. By comparing the video footage to the log, we later checked if users had pressed the button with their name or the button that said ‘I’m not listed’ if this was the case.

7.3.3 Results

The test session and the questionnaire were completed by all 30 participants. During the usage period, the system was used 57 times by these registered users, with a total usage time of 163 minutes. Table 7.2 presents the descriptive statistics. It shows that the maximum number of sessions by a particular user was 9 and the longest time it was used, was 16 minutes. Two participants did not use the system at all.

Since we measured usage both in ‘number of sessions’ and ‘minutes of use’ we had two possibilities to link Intention to Use to actual use. However, there were users that had a range of short sessions, logging out after each menu item was used and logging in again to do the next thing. Others used it for a longer period without logging out. This means that the number of sessions would not give an accurate indication. So we chose not to use the number of times, but the amount of minutes per user (which often would be an accumulation of several times of use) as usage indication.

	Minimum	Maximum	Mean	Std. Deviation
Anxiety	2.75	5.00	3.958	.606
Attitude	1.50	5.00	3.800	.943
Facilitating Conditions	1.50	5.00	3.317	1.038
Intention to Use	1.00	4.75	3.225	1.371
Perceived Adaptivity	1.67	5.00	3.411	1.027
Perceived Enjoyment	1.60	5.00	3.667	.844
Perceived Ease of Use	1.00	5.00	3.467	1.133
Perceived Sociability	1.25	4.75	3.350	.872
Perceived Usefulness	1.00	5.00	2.989	1.056
Social Influence	2.00	5.00	3.300	.6900
Social Presence	1.00	4.75	2.600	1.317
Trust	1.00	5.00	3.567	1.128
Number of sessions	0	9	1.90	2.339
Minutes	0	16	5.43	4.606

Table 7.2. Descriptive Statistics Experiment 5

Again, we calculated Cronbach's Alpha to test the reliability of the constructs. As Table 7.3 shows, the constructs were reliable. To obtain these scores we had to omit statements 10 (FC3), 21 (PENJ5), 22 (PEOU1), 37 (SI3) and 45 (Trust3). These questions have also been omitted in previous experiments. In Chapter 8 we will draw conclusions on permanently deleting statements from the list.

Construct	*Alpha*	*Construct*	*Alpha*
Anxiety	*.754*	*Perceived Ease of Use*	*.820*
Attitude	*.801*	*Perceived Sociability*	*.786*
Facilitating Conditions	*.706*	*Perceived Usefulness*	*.787*
Intention to Use	*.947*	*Social Influence*	*.793*
Perceived Adaptivity	*.834*	*Social Presence*	*.866*
Perceived Enjoyment	*.836*	*Trust*	*.802*

Table 7.3. Cronbach's Alpha Experiment 5

To test the model validation hypotheses, we performed a regression analysis on the scores. The results are shown in Table 7.4.

The high t-score for Hypothesis 1(a) implies that for this setting with this robot, the assumption that Intention to Use predicts actual use has clearly been confirmed. This is despite the low R^2 value for this hypothesis, which is not unusual since there can be many external influences on actual use. Social Influence and Facilitating Conditions however, while reliable constructs in this case, have no proven influence on Usage. Furthermore, the R^2 value for Hypothesis 2 is .63, which is satisfactory. A projection of the results of the regression analysis on the visualization of our model shows that some constructs appear to be without influence in this case (see Figure 7.3). Remarkable is the high contribution of Perceived Adaptiveness on Perceived Usefulness, implying that a system that is perceived as more adaptive/adaptable is perceived to be useful. However, Perceived Usefulness is for this robot in this setting not a determining influence on Intention to Use.

Hypothesis		Independent	Dependent	Beta	t	Sig.
H1	(a)	ITU		.671	4.603**	.000
	(b)	SI	Usage(min)	.133	.937	.357
	(c)	FC		-.276	-1.909	.067
		Model: R2=.49; F=8.243; df=3,26; P=.001				
H2	(a)	PU		.094	.510	.615
	(b)	PEOU		.545	3.373**	.003
	(c)	ATT	ITU	.437	1.946*	.049
	(d)	PENJ		-.078	-.347	.731
	(e)	SI		-.225	-1.491	.149
	(f)	Trust		.044	.223	.826
		Model: R^2=.63; F=5.279; df=6,23; P=.002				
H3	(a)	ANX		.293	2.120*	.043
	(b)	ATT	PU	.029	.209	.836
	(c)	PAD		.339	2.141*	.041
	(d)	PEOU		.324	2.177*	.038
		Model: R^2=.59; F=9.814; df=4,25; P=.000				
H4	(a)	ANX	PEOU	.268	1.854	.075
	(b)	PENJ		.572	3.962**	.000
		Model: R^2=.54; F=15.809; df=2,27; P=.000				
H5	(a)	PS	PENJ	.361	2.144*	.041
	(b)	SP		.411	2.446*	.021
		Model: R^2=.46; F=11.479; df=2,27; P=.000				
H6		Trust	PS	.418	2.435*	.022
		Model: R^2=.18; F=5.929; df=1,28; P=.022				
H7		PS	SP	.540	3.399**	.002
		Model: R^2=.29; F=11.551; df=1,28; P=.002				

Table 7.4. Regression scores on hypotheses Experiment 5

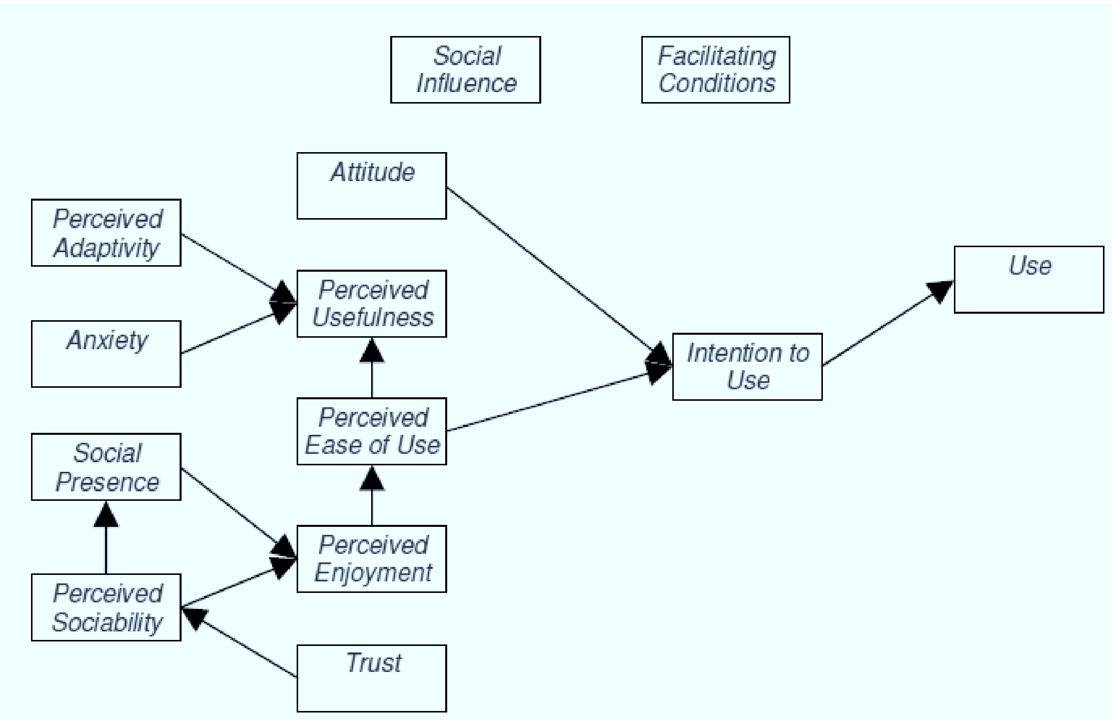


Figure 7.3. Confirmed hypotheses for Experiment 5

7.4 Experiment 6: actual use in a private setting

For Experiment 6, we used Steffie, a service type screen agent with companion type features. Steffie (see Figure 7.4) is designed in Flash and developed as a part of a website (www steffie.nl) where she features as a talking guide for older adults, explaining the internet, e-mail, health insurance, cash dispensers and railway ticket machines. The user communicates with her by clicking buttons that are used for choosing subjects, to let her continue or to let her repeat. Steffie has been developed by a consortium of commercial and non-commercial partners, as a part of a project to facilitate the use of the internet by older adults.

We used an offline version of the application, kindly provided to us by the developers. We used this version on the pc's of participants at their homes. We added an entrance page on which there were the names of possible users (Figure 7.5). Usually with an elderly couple this would be both names and an entrance for 'someone else' for a possible visitor. If the user chose a name, it was recorded in a log file and if the user ended the session, it wrote the ending time in the log file. Also, if the user did not use the application for 90 seconds, it closed and wrote the time in the log file.

Figure 7.4. Screenshot of Steffie

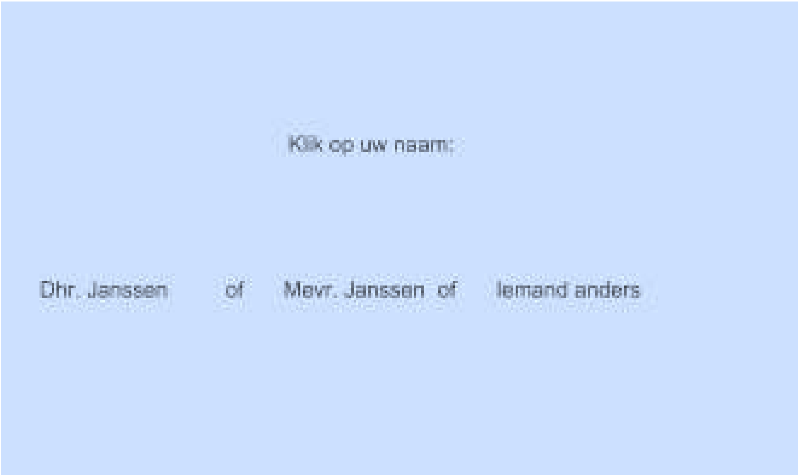


Figure 7.5. Login screen for the Steffie Application

7.4.1 Participants

Participants were 30 elderly users who owned a PC. Their age ranged from 65 to 89 and they were all living independently. Of the 30 participants, 14 were female, 16 were male. Table 7.5 shows their descriptive statistics. Note that the mean score for Experience is exceptionally high compared to the previous experiments.

	Minimum	Maximum	Mean	Std. Deviation
Age	65	89	72.33	6.41
Experience	1.00	5.00	3.48	1.29
Education	2	10	6.13	2.58

Table 7.5. Age, education and computer experience Experiment 6

7.4.2 Procedure

The participants were visited by a researcher who installed the Steffie application on their PC. Subsequently they were to try out the application for a minimum of two and a maximum of three minutes. After this they were to fill out our questionnaire. After ten days, the researcher returned, copied the log file and deleted the application from the pc.

7.4.3 Results

All 30 participants completed their forms and we were able to collect usage data for each of them (although one participant did not use the application at all). Desciptive statistics are shows in Table 7.6.

Again, we calculated Cronbach's Alpha for the used constructs to see if they were consistent. As Table 7.5 shows, the constructs had high scores and can be considered reliable. Again, to obtain these scores we had to omit statements 10

(FC3), 21 (PENJ5), 22 (PEOU1), 37 (SI3) and 45 (Trust3) which were the same as in our previous experiment.

	Minimum	Maximum	Mean	Std. Dev.
Anxiety	*2.00*	*5.00*	*4.233*	*.731*
Attitude	*2.00*	*5.00*	*3.744*	*.815*
Intention to Use	*1.00*	*5.00*	*4.050*	*.981*
Facilitating Conditions	*1.00*	*5.00*	*3.900*	*.875*
Perceived Adaptivity	*2.00*	*5.00*	*3.767*	*.541*
Perceived Enjoyment	*2.67*	*4.67*	*3.793*	*.519*
Perceived Ease of Use	*2.80*	*4.60*	*3.887*	*.596*
Perceived Sociability	*2.60*	*4.80*	*3.633*	*.721*
Perceived Usefulness	*1.75*	*4.75*	*3.711*	*.715*
Social Influence	*2.00*	*4.67*	*3.350*	*.604*
Social Presence	*2.00*	*4.00*	*2.727*	*.686*
Trust	*2.00*	*4.20*	*3.717*	*.639*
Times	*0*	*10*	*3.20*	*2.441*
Minutes	*0*	*134*	*33.93*	*33.686*

Table 7.6. Descriptive Statistics Experiment 6

Construct	Alpha	Construct	Alpha
Anxiety	*.812*	*Perceived Ease of Use*	*.726*
Attitude	*.869*	*Perceived Sociability*	*.878*
Intention to Use	*.707*	*Perceived Usefulness*	*.865*
Facilitating Conditions	*.948*	*Social Influence*	*.794*
Perceived Adaptivity	*.709*	*Social Presence*	*.816*
Perceived Enjoyment	*.774*	*Trust*	*.732*

Table 7.7. Cronbach's Alpha Experiment 6

To test our model validation hypotheses, we performed a regression analysis. As Table 7.8 shows, hypothesis 1 could be fully confirmed (with a high t-score for H1(a)). However, the influences of Social Influence and Facilitating Conditions are negative. Regarding Social Influence this indicates that the more users are concerned by what others think of their use of the system, the less they actually used it. The reason for this pattern might be, that the application is actually developed for people who need help on using common applications like the internet, railway ticket machines and cash dispensers. Participants who saw the benefits of the application, may have reasoned that needing this explanation does not leave a good impression. In this sense, there may have been a strong influence of *stigmatization* – which has earlier been identified as an influential factor in research on assistive robotics (Forlizzi et al. 2004).

Also the negative influence of Facilitating Conditions demands an explanation. We have to consider that the application was explicitly developed for older adults with little computer experience. Users with less computer experience can thus be expected to have a higher score on Intention to Use. However, because of their little experience, they may not have felt confident on being able to use it properly. The participants who had more computer experience may have noticed that they would not need any help to use it properly, but because of their experience they would have needed it less.

Hypothesis		Independent	Dependent	Beta	t	Sig
H1	(a)	ITU	Usage(min)	.945	5.700**	.000
	(b)	SI		-.420	-2.553*	.017
	(c)	FC		-.327	-2.576*	.016
		Model: R^2=.59; F=12.508; df=3,26; P=.000				
H2	(a)	PU	ITU	.616	3.301**	.003
	(b)	PEOU		-.114	-.794	.435
	(c)	ATT		.491	2.228*	.036
	(d)	PENJ		-.125	-.668	.510
	(e)	SI		.015	.105	.917
	(f)	Trust		.011	.067	.947
		Model: R^2=.79; F=15.600; df=6,23; P=.000				
H3	(a)	ANX	PU	.310	2.589*	.016
	(b)	ATT		.302	1.982	.059
	(c)	PAD		.337	2.444*	.022
	(d)	PEOU		.179	1.203	.240
		Model: R^2=.73; F=17.400; df=4,25; P=.000				
H4	(a)	ANX	PEOU	.268	1.854	.075
	(b)	PENJ		.572	3.962**	.000
		Model: R^2=.51; F=13.923; df=2,27; P=.000				
H5		PS	PENJ	.686	4.259**	.000
		SP		.097	.600	.554
		Model: R^2=.56; F=17.190; df=2,27; P=.000				
H6		Trust	PS	.608	4.057**	.000
		Model: R^2=.37; F=16.456; df=1,28; P=.000				
H7		PS	SP	.609	4.067**	.000
		Model: R2=.37; F=16.544; df=1,28; P=.000				

Table 7.8. Regression scores on model validation hypotheses for Experiment 6

As Table 7.9 shows, a correlation analysis supports this interpretation with (a) a negative correlation between Experience and Minutes and (b) a strong positive correlation between Experience and Facilitating Conditions. This indicates that the more experienced users (a) used the application less and (b) they were more confident on being able to use the system properly.

		Experience	FC
Minutes	Pearson Correlation	-,423*	-,215
	Sig. (2-tailed)	,020	,254
FC	Pearson Correlation	,487**	1
	Sig. (2-tailed)	,006	

Table 7.9. Correlation score for Computer Experience, Facilitating Conditions and usage in minutes for Experiment 6

Hypothesis 2 could only partly be confirmed: only Perceived Usefulness and Attitude could be established as determining influences on Intention to Use. Hypotheses 3, 5, 6 and 7 could be fully confirmed. For Hypothesis 4, only (b) could be confirmed, leaving Perceived Enjoyment as the only determining influence on Perceived Ease of Use.

We conclude that Intention to Use indeed is a predictor of actual use for this technology, and Perceived Usefulness is the dominating influence on Intention to Use. Where it could be expected that the application's ease of use would be the main reason for users to be enthusiastic about it, it is the usefulness that seems to make Steffie attractive to her users. Perceived Ease of Use does contribute, but only through its influence on Perceived Usefulness, together with Anxiety and Perceived Adaptivity. Other results confirm what we found in our previous experiments: Social Presence is influenced by social abilities that are perceived; Social Presence is influencing Perceived Enjoyment. But Perceived Enjoyment in this case does not directly influence Intention to Use.

The visualization of the regression results (Figure 7.6) shows that the determining influence on Intention to Use is limited to Attitude and Perceived Usefulness. This suggests that this application is primarily used for its functionality and not so much for the enjoyment it provides. If we take into consideration that Steffie is very much a utilitarian application, this is not surprising. The iCat application - which did not have Perceived Usefulness as a major determining influence - had functionalities that could be experienced less as primarily useful and more as enjoyment related. This was even more so in Experiment 4, in which Perceived Enjoyment was found to directly determine Intention to Use (Figure 5.4).

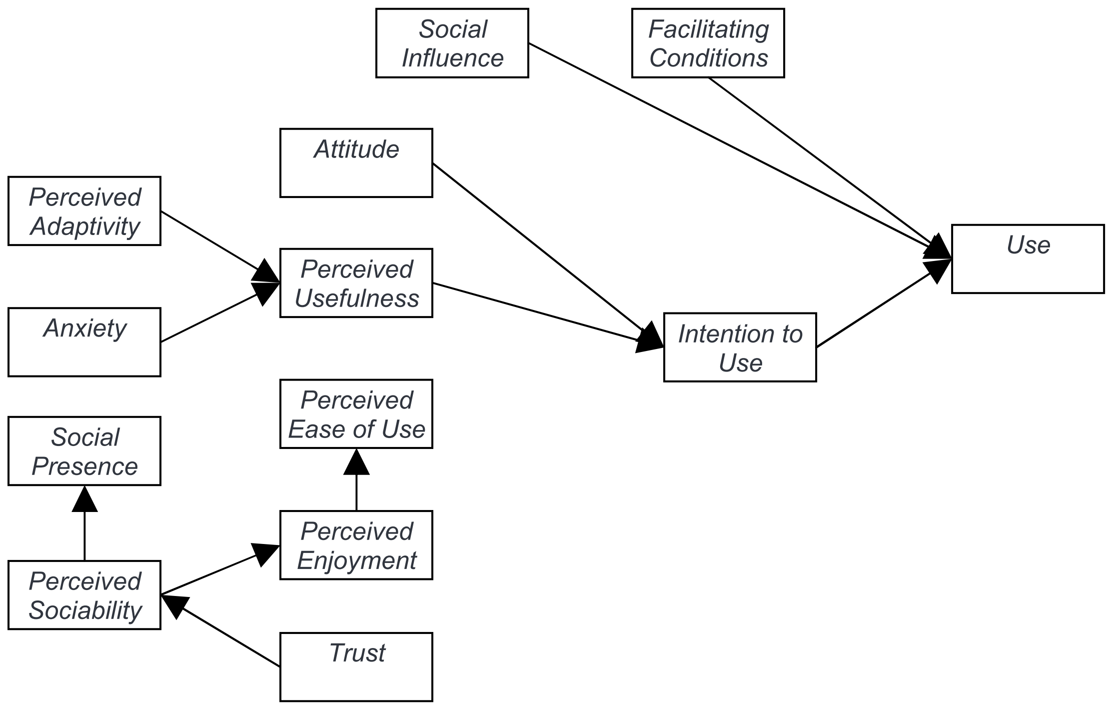


Figure 7.6. Confirmed hypotheses for Experiment 6

Furthermore, as with the RoboCare movies in Experiment 4, Social Presence is not of influence. It could be that it is experienced less intensely, because it is indirect (through a screen).

Another difference is that there are no constructs without confirmed interrelations, as there were Social Influence and Facilitating Conditions in the previous experiment. The influence of these particular constructs however, is a negative one, which – as we discussed earlier – could have to do with the stigmatizing type of application.

7.5 Combined results

Since the questionnaire was the same for both systems, we can combine the results. This will give us an indication of relationships that can be generalized for different assistive social robots, used in different settings. An overview of descriptive statistics is given in Table 7.10. It shows a high mean score with a relative low standard deviation for Anxiety, indicating that in general, participants felt comfortable with these systems. This is however to be expected: participants were older adults that volunteered for the experiment and were thus likely to be curious about the new technology - older adults that would feel much anxiety on new technology were not very likely to volunteer.

We carried out an ANOVA to determine whether user responses differed between the two systems. As Table 7.10 shows, there are differences concerning Facilitating Conditions and Intention to Use: both were much higher in Experiment 6 (screen agent). We have to be careful explaining this, since the two systems differed in more than one aspect. However, regarding Intention to Use, we may take into consideration that in Experiment 6 the participants could use the application on their own computer at home, so we can state that the setting was more inviting to use it. Regarding the higher score on Facilitating Conditions, we assume that the participants in Experiment 6 were generally confident that the application was simple enough. As Table 7.9 showed, this is even more so for the more experienced participants.

Dependent Variable	**Type III Sum of Squares**	**Mean Square**	**F**	**Sig.**
Anxiety	.704	.704	1.726	.194
Attitude	.046	.046	.064	.801
Intention to Use	5.104	5.104	5.540*	.022
Facilitating Conditions	10.209	10.209	7.180*	.010
Perceived Adaptivity	1.896	1.896	2.815	.099
Perceived Enjoyment	.241	.241	.490	.487
Perceived Ease of Use	.600	.600	1.052	.309
Perceived Sociability	1.204	1.204	1.879	.176
Perceived Usefulness	1.956	1.956	2.614	.111
Social Influence	.017	.017	.045	.832
Social Presence	.241	.241	.218	.642
Trust	.338	.338	.402	.529

Table 7.10. ANOVA combined results, comparing the used systems

On these combined results we performed a regression analysis, testing the model validation hypotheses. The results presented in Table 7.12 show that only

Intention to Use remains as a determining influence on Usage. The determining influences on Intention to Use are Perceived Usefulness, Perceived Ease of Use and Attitude. The strongest influences on Perceived Usefulness is Perceived Adaptiveness, but also Anxiety and Perceived Ease of Use are determining influences. Figure 7.7 shows a visualization of this analysis.

Construct	Minimum	Maximum	Mean	Std. Dev.
Anxiety	2.00	5.00	4.342	.643
Attitude	1.33	5.00	3.683	.843
Facilitating Conditions	1.00	5.00	3.608	.996
Perceived Adaptivity	1.00	5.00	3.638	1.253
Perceived Enjoyment	1.67	5.00	3.589	.833
Anxiety	1.60	5.00	3.730	.698
Perceived Ease of Use	2.00	5.00	3.787	.755
Perceived Sociability	1.25	4.75	3.492	.806
Perceived Usefulness	2.00	5.00	3.531	.877
Social Influence	2.00	5.00	3.367	.603
Social Presence	1.00	4.75	2.663	1.043
Trust	1.00	5.00	3.642	.912

Table 7.11. Descriptive Statistics combined results

Hypothesis	Independent	Dependent	Beta	t	Sig
H1 (a)	ITU	Usage(min)	.584	4.718**	.000
(b)	SI		-.096	-.803	.426
(c)	FC		-.126	-1.072	.288
Model: R^2=.29; F=7.600; df=3,26; P=.000					
H2 (a)	PU	ITU	.546	4.410**	.000
(b)	PEOU		.252	2.320*	.024
(c)	ATT		.457	3.498**	.001
(d)	PENJ		-.183	-1.375	.175
(e)	SI		-.130	-1.362	.179
(f)	Trust		-.044	-.372	.712
Model: R^2=.67; F=18.420; df=6,23; P=.000					
H3 (a)	ANX	PU	.214	2.725*	.009
(b)	ATT		.059	.655	.516
(c)	PAD		.669	8.039**	.000
(d)	PEOU		.174	2.292*	.026
Model: R^2=.76; F=43.398; df=4,25; P=.000					
H4 (a)	ANX	PEOU	.175	1.348	.183
(b)	PENJ		.335	2.582*	.012
Model: R^2=.19; F=6.788; df=2,27; P=.002					
H5 (a)	PS	PENJ	.449	3.936**	.000
(b)	SP		.338	2.959**	.004
Model: R^2=.48; F=26.560; df=2,27; P=.000					
H6	Trust	PS	.477	4.134**	.000
Model: R^2=.23; F=17.090; df=1,28; P=.000					
H7	PS	SP	.550	5.014**	.000
Model: R^2=.30; F=25.142; df=1,28; P=.000					

Table 7.12. Regression scores on hypotheses for the combined results

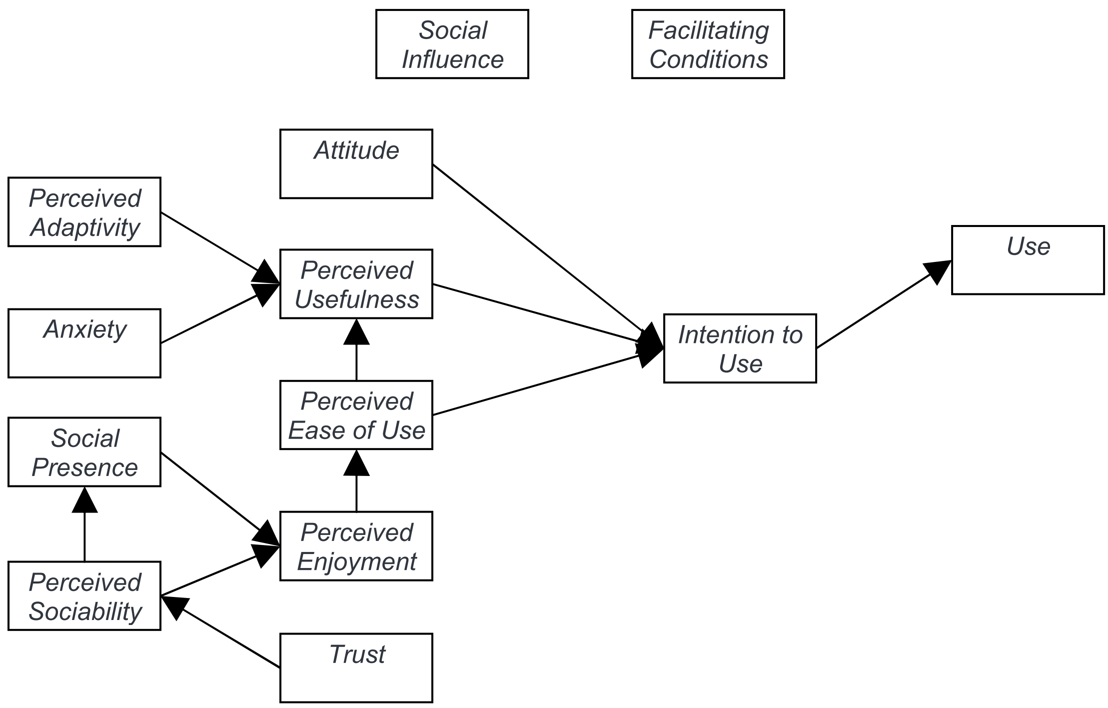


Figure 7.7. Confirmed hypotheses for the combined results

7.6 Moderating factors

To test our hypotheses on moderating factors (set up in section 4.4), we again performed a Chow's test on each set of results. Table 7.13 shows the outcome for the iCat study, in which we only could test the first two hypotheses and hypothesis 5a, since the other hypotheses concerned relations that were not confirmed in the regression analysis. As the table shows, we could not confirm any moderating effect.

Hypothesis	**Factor*Variable**	**Dependent**	**F**	**Sig.**
MFH1a	Gndr*PU	ITU	.297	.591
MFH1b	Age*PU	ITU	1.192	.379
MFH2a	Gndr*PEOU	ITU	.062	.804
MFH2b	Age*PEOU	ITU	1.299	.321
MFH2c	EXP*PEOU	ITU	.478	.752
MFH5a	EDU*ANX	PU	1.815	.131

Table 7.13. Chow's test on moderating factor hypotheses for Experiment 5

Table 7.14 shows the outcome for the Chow's test using the results of the experiment with Steffie. Since hypotheses 2 and 3 concerned influences that were not confirmed in our regression analysis, we could only test hypotheses 1, 4 and 5. The results show a significant moderating effect for Experience concerning hypothesis 4b (Facilitating Conditions determining Usage) and for Education, concerning hypothesis 5a and 5b (Anxiety determining Perceived Usefulness and Perceived Ease of Use).

Hypothesis	Variable*Factor	Dependent	F	Sig.
MFH1a	Gndr*PU	ITU	.543	.380
MFH1b	Age*PU	ITU	.913	.571
MFH4a	Age*FC	Use	3.472	.128
MFH4b	Exp*FC	Use	2.982*	.025
MFH5a	EDU*ANX	PU	3.835*	.008
MFH5b	EDU*ANX	PEOU	3.099*	.021

Table 7.14. Chow's test on moderating factor hypotheses for Experiment 6

Table 7.15 shows that the combined results do not show any significant moderating influences. Again, we could only test the hypotheses concerning influences that were confirmed in the regression analysis, which were in this case 1, 2 and 5a.

Hypothesis	Variable*Factor	Dependent	F	Sig.
MFH1a	Gndr*PU	ITU	.562	.781
MFH1b	Age*PU	ITU	1.441	.266
MFH2a	Gndr*PEOU	ITU	.171	.680
MFH2b	Age*PEOU	ITU	1.195	.310
MFH2c	EXP*PEOU	ITU	.766	.676
MFH5a	EDU*ANX	PU	2.082	.062

Table 7.15. Chow's test on moderating factor hypotheses for the combined results

7.7 Conclusions

We have tested our new model in two different usage settings, with two very different systems. In both cases actual usage was predicted by the indicated Intention to Use the system. This is a clear result. For the different hypotheses, the conclusions are more complex. We will discuss them one by one.

H1 Use is determined by (a) Intention to Use and influenced by (b) Social Influence and (c) Facilitating Conditions.

Only data from the screen agent experiment confirmed the influence of (b) Social Influence and (c) Facilitating Conditions - and for both it is negative. Social Influence for this user group is apparently potentially relevant, but has to be interpreted slightly different from its original meaning. In the UTAUT model it refers to appearing fashionable or generally impressing peers and superiors, while with technology developed for older adults it can be referring more specifically to stigmatizing aspects. For technology that has not explicitly been developed for older adults, like the iCat, this influence could thus be less relevant for this user group.

Also the concept of Facilitating Conditions has to be interpreted differently than in the UTAUT model. In a working environment it can be related to the availability (or non availability) of a helpdesk, or a manual. In this non working context, where the users do not need to 'perform' with the tested technology, the replies to the statements of Facilitating Conditions can be related to experience and feeling confident. This could explain the fact that this construct has no determining influence in Experiment 5, where neither help nor experience were required to appropriately use the application. In Experiment 6 however, Experience is very influential and modifies the response to the Facilitating Conditions statements. And as it concerned an application developed to help the less experienced, a higher score on Experience corresponded both with a higher score on Facilitating Conditions and a lower score on actual use (Table 7.9).

An analysis of descriptive statistics (Table 7.16) for both systems confirms this analysis, showing a relatively high score on Computer Experience and on Facilitating Conditions for the Users of Steffie.

	iCat	***St.D.***	***Steffie***	***St.D.***
Computer Experience	*2.133*	*1.629*	*3.483*	*1.290*
Facilitating Conditions	*3.317*	*1.038*	*3.900*	*.875*
Social Influence	*3.300*	*.6900*	*3.350*	*.604*

Table 7.16. Selection of mean scores for Experiments 5 and 6

Also our finding in the analysis of moderating factors that Experience has a moderating effect on Facilitating Conditions determining Usage confirm the explanation we gave earlier for the negative effect of Facilitating Conditions.

Furthermore, with the iCat with touch screen application, Experience could have had no effect, because users could have found it so easy to use that Facilitating Conditions were irrelevant.

H2 Intention to Use is determined by (a) Perceived Usefulness, (b) Perceived Ease of Use, (c) Attitude, (d) Perceived Enjoyment, (e)Social Influence and (f)Trust.

The influences of (d) Perceived Enjoyment, (e) Social Influence and (f) Trust could not be confirmed. The influence of Perceived Ease of Use was only confirmed for the iCat with touch screen, the influence of Perceived Usefulness only for the screen agent. The influence of Attitude could be confirmed for both.

These findings show how the model reflects that influences on Intention to Use are varying, depending on robot characteristics and tasks. For a screen agent being developed to provide guidance, Perceived Usefulness is characteristically determining (it will mainly be used for its usefulness) and for a robot that is used by a touch screen, Perceived Ease of Use seems essential. Still, in the latter case, it could very well have been that the touch screen that was used contributed to the experienced Ease of Use, although the questions were explicitly about the iCat, and not on the system. It would be interesting to see, if Perceived Ease of Use is a determining influence in settings without a touch screen.

H3 Perceived Usefulness is influenced by (a) Perceived Ease of Use (b) Attitude, (c) Perceived Adaptivity and (d) Anxiety

For both experiments , part (c) and (d) of this hypothesis could be confirmed and for Experiment 5 also part (a). Attitude was not among the significant influences on Perceived Usefulness in both experiments, possibly because its influence has been suppressed by the strength of the other influences.

H4 Perceived Ease of Use is influenced by (a) Anxiety and (b) Perceived Enjoyment

The influence of Perceived Enjoyment could be confirmed for both systems. Anxiety was not a significant determining influence in Experiment 5: the enjoyment was apparently so strong that it was not suppressed by it.

H5 Perceived Enjoyment is influenced by (a) Social Presence and (b) Perceived Sociability
H6 Perceived Sociability is influenced by Trust
H7 Social Presence is influenced by Perceived Sociability

These hypotheses could be confirmed for both systems, except for H5(a): Social Presence was not determining Perceived Enjoyment in Experiment 6, which may be due to the less direct experience of a social entity with a screen agent.

Concerning the moderating influences, we could only confirm the following:

- Experience has a moderating effect on Facilitating Conditions determining Usage: for users with a higher score on Experience, Facilitating Conditions are less determining than for users with more experience.
- Education has an effect on the influence of Anxiety on both Perceived Usefulness and Perceived Ease of Use: the higher user scored on education, the less Anxiety was of influence on both constructs.

These effects were only found for the results on Steffie.

Figure 7.8 shows a visualization of the results for both systems and the combined results.

Regarding the R^2 values on Intention to Use (.63 for the iCat experiment, .79 for the Steffie experiment) we can establish that they are relatively high when compared to our first two experiments. The higher value (.16 difference) for Steffie could be explained by the setting (at home), the system (on screen embodiment) or the tasks. In the next chapter, where we overview the results of all experiments, we will reflect more on this.

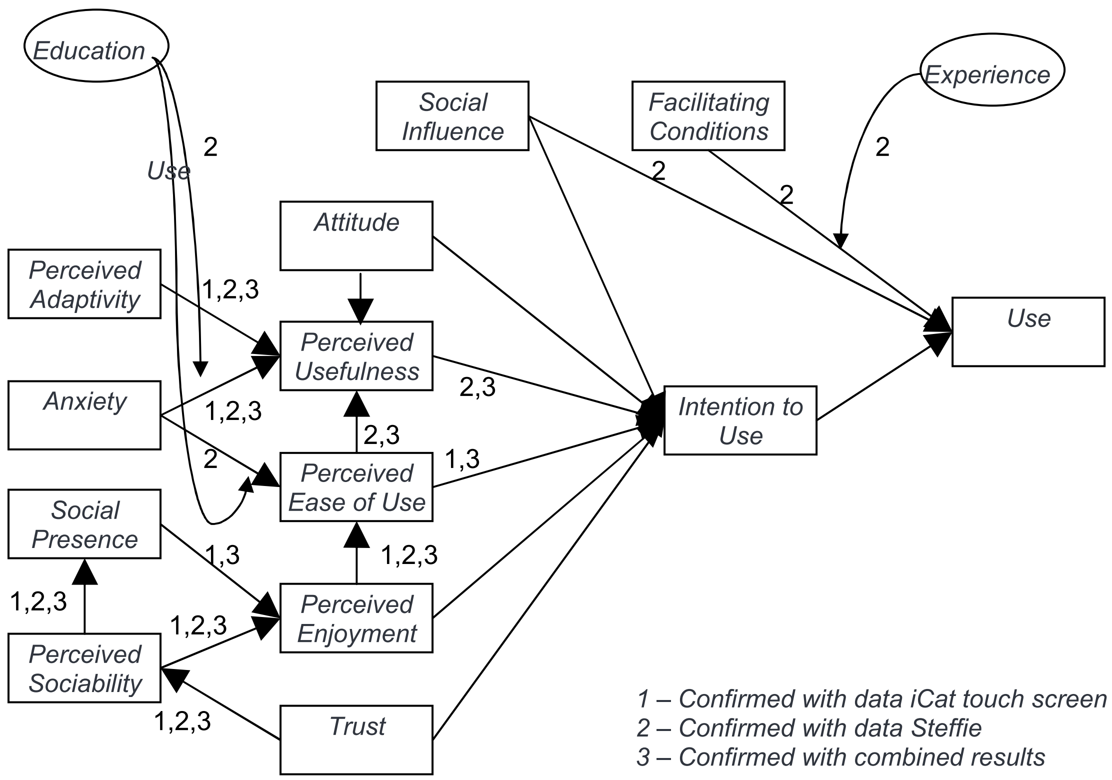


Figure 7.8. Confirmed hypotheses both systems and combined results

8. Developing a new model II

Parts of this chapter have been published earlier in (Heerink et al. 2010b)

8.1 Introduction

In this chapter, we will first overview and compare the results of the six experiments we did (two with the initial model and four with the new model). Next, we will establish whether combining the results of the four experiments for which the same questionnaire was used (see Table 8.1) gives us additional findings on construct interrelations.

The possibility to combine results also means we have a relatively high number of cases, which gives us the opportunity to carry out a path analysis using structural equation modeling (see Section 2.4.4), which also may give us additional insight concerning construct interrelations and the strength of our model as a whole. This statistical technique is similar to a regression analysis, but can be applied to an entire model instead of testing hypotheses concerning parts of it as we did in the previous chapters. Moreover, it can be used to explore the possibility of additional construct relations, which is what we will do in section 8.4.

The result of these procedures is the establishment of our final and fine-tuned model and methodology.

8.2 Evaluating experimental results

Since the beginning of our study, we carried out six experiments: two to evaluate the applicability of the UTAUT model (Chapter 3), two to justify the addition of unprecedented constructs (social constructs and adaptivity) and two to validate our newly developed model in a setting that included the gathering of usage data (public and private). Table 8.1 lists these experiments, also showing the R^2 outcomes. The table shows that the R^2 values are generally higher in the experiments 3 to 6 in which our newly developed methodology was used (average R^2 for these is .72, while it was .47 for the first two experiments). Nevertheless,

R^2 values are still varying from .62 up to .88. This means we can state that our new model performs better, but it does not show an equal predictive power in different conditions.

Furthermore, R^2 values are higher for the screen agents. In our first two experiments, we found the model performed better with the results on screen agent Annie than with the results on iCat, and in the experiments in which the new model was used we found the R^2 value for Steffie to be higher than the (average) R^2 values concerning the other systems.

	Experiment	System	N	Chapter	R^2 ITU
UTAUT	**1** iCat social abilities	iCat robot WOZ in 2 conditions	36	3	.37
	more social		17		.28
	less social		19		.45
	2 Annie social abilities	Screen agent WOZ	36	3	.59
	more social		18		.50
	less social		18		.65
	Average 1 and 2				*.47*
	Combining 1 en 2		72		.45
	more social		35		.34
	less social		37		.52
New model	**3** iCat social abilities 2	iCat robot WOZ	40	5	.70
	more social		20		.72
	less social		20		.71
	4 Robocare adaptivity	Robocare robot videos in four conditions	88	6	.68
	neutral		22		.69
	adaptable		21		.88
	adaptive + user control		23		.69
	adaptive - user control		22		.62
	5 iCat public usage	'autonomous' iCat robot with touch screen	30	7	.63
	6 Steffie private usage	'autonomous' screen agent	30	7	.79
	Combining 5 en 6		60		.67
	Average 3-6				*.71*

Table 8.1. Listing of experiments

We have some difficulty explaining this, because the screen agents differed in more than one aspect from the other systems: they were humanoid and they were presented on a screen. Moreover, the setting and the functionalities for Steffie were different from those for the other systems and (as discussed in the previous chapter), it was also used by participants with more experience. It could be that the processes described by the model are more applicable to more experienced users.

If the screen presence is causing this difference, it could be that adding an influence related to the physical embodiment of the robot would lead to a better performance of the model. However, it could also be that interaction through a computer screen is a more common experience with a more common effect on the users. The experience of a futuristic technology like a robot may have a different effect on older individuals, thus causing the data to be less homogeneous. An indication for this effect could be given by lower values for standard deviations for the screen agent. As we will see later in this section, Table 8.2 shows this is indeed the case.

Both Figure 8.1 and Table 8.2 give an overview of the results of the experiments with the new model, the first showing the different confirmed construct interrelations and the second comparing the mean scores on the constructs.

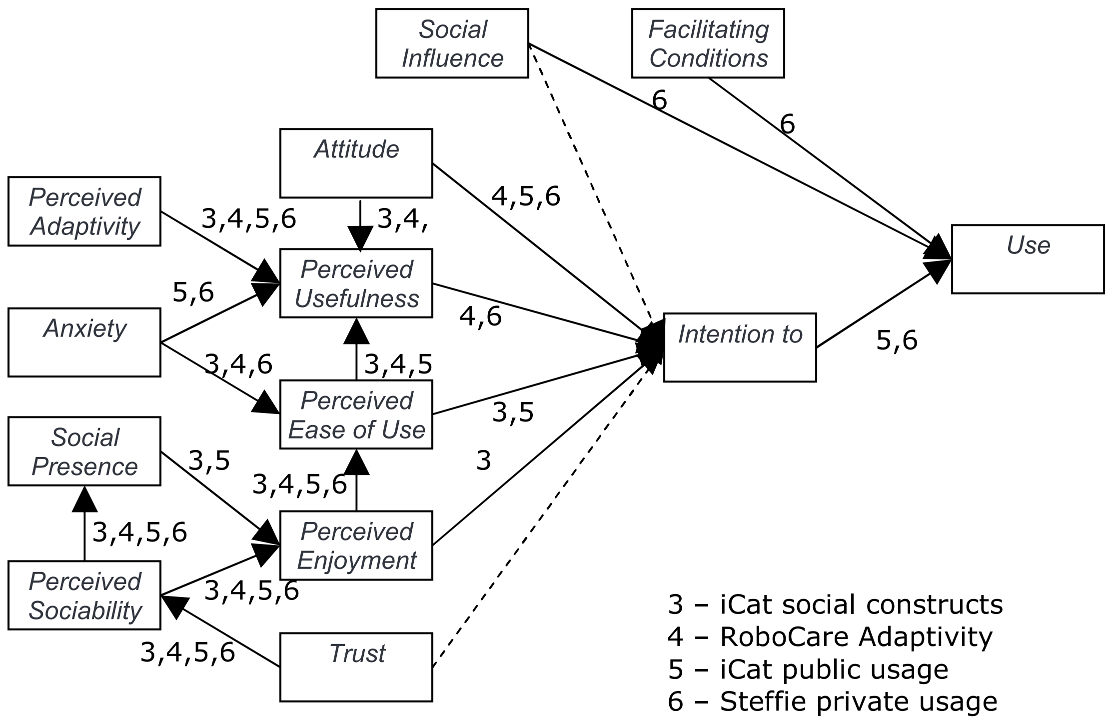


Figure 8.1. Confirmed construct interrelations for the new model experiments. The dotted lines represent hypothesized construct interrelations that have not been confirmed by any of the experiments.

First of all, Figure 8.1 shows how the confirmed construct interrelations (straight lines) vary: some are confirmed in one experiment but not in another. These differences reflect the differences between the used systems, chosen settings and performed tasks. They demonstrate how the model can be used to explain the impact of changes in a system, a setting or a task on the perception of the system by the users. We applied this in the experiments where we compared different conditions of the same system and in Chapter 7, where we compared the two usage experiments. The latter demonstrates that a comparative analysis can lead to new insights even if systems differ in more than one aspect.

Analyzing the confirmed construct interrelations in Figure 8.1, we can state that the predominant influences on Intention to Use are Perceived Usefulness and Attitude for both Steffie (experiment 6) and Robocare (Experiment 4). As we concluded earlier, for these systems that are more utilitarian, functionality is crucial for acceptance. For the experiments with the more ‘hedonic’ iCat (experiments 3 and 5), Perceived Ease of Use and Perceived Enjoyment are the main determining influences on acceptance. This means the predominant influences on Intention to Use can be seen as a reflection of the hedonic or utilitarian nature of the system.

In Table 8.2 we compare the descriptive statistics of the new model experiments, which can provide additional insight on what the used systems do and do not have in common.

	Exp. 3		***Exp. 4***		***Exp. 5***		***Exp. 6***	
	Mean	***St.D.***	***Mean***	***St.D.***	***Mean***	***St.D.***	***Mean***	***St.D***
Anxiety	*2.341*	*1.053*	*3.671*	*.734*	*3.958*	*.606*	*4.233*	*.731*
Attitude	*3.383*	*1.023*	*3.167*	*.922*	*3.800*	*.943*	*3.744*	*.815*
Facilitating Conditions	-	-	*3.477*	*.739*	*3.317*	*1.038*	*3.900*	*.875*
Intention to Use	*3.288*	*1.018*	*3.402*	*1.052*	*3.225*	*1.371*	*4.050*	*.981*
Perceived Adaptivity	*2.842*	*.955*	*3.492*	*.645*	*3.411*	*1.027*	*3.767*	*.541*
Perceived Enjoyment	*3.900*	*.663*	*2.955*	*.860*	*3.667*	*.844*	*3.793*	*.519*
Perceived Ease of Use	*3.525*	*1.054*	*3.559*	*.714*	*3.467*	*1.133*	*3.887*	*.596*
Perceived Sociability	*3.750*	*.737*	*2.696*	*.645*	*3.350*	*.872*	*3.633*	*.721*
Perceived Usefulness	*3.200*	*1.088*	*3.633*	*.893*	*2.989*	*1.056*	*3.711*	*.715*
Social Influence	*3.206*	*.882*	*3.205*	*.730*	*3.300*	*.6900*	*3.350*	*.604*
Social Presence	*3.600*	*.914*	*2.830*	*.533*	*2.600*	*1.317*	*2.727*	*.686*
Trust	*2.630*	*.879*	*3.602*	*.898*	*3.567*	*1.128*	*3.717*	*.639*

Table 8.2. Means and standard deviation for the new model experiments

If we continue to compare the means in Table 8.2, we see a relatively high score on Social Presence for Experiment 3, which coincides with a relatively low score on Trust, Perceived Adaptivity and Anxiety (which actually means that Anxiety is higher). The high score for Social presence can be easily explained: this was the only experiment in which participants had a full conversation with the system. In the other experiments the interaction was less direct (through touch screen or mouse clicks) or there was even no interaction by the participants (in the RoboCare movies) which decreased the sense of presence. The fact that the participants had probably hardly any experience with this voice driven interaction, could have contributed to the Anxiety score. However, future research will have to confirm that a new form of interaction will indeed cause elderly users to feel more insecure.

The scores on Trust and Perceived Adaptivity can be explained by the relatively limited functionality, which gave the system very little opportunity to gain trust or demonstrate adaptivity.

Returning to to Figure 8.1, we furthermore conclude that the influence of Social Presence on Perceived Enjoyment has only been significant for the two experiments with a ‘real three dimensional’ robot (iCat in experiment 3 and 5). In the experiments where we used a screen agent (6) or a video (4), the participants were not directly confronted with a robot as a physical (three dimensional) entity, which explains that their sense of presence was less strong –

which we already concluded from the mean scores - and less (or even not at all) of influence on their enjoyment. However, some influences turn out to be general:

- Trust determining Perceived Sociability;
- Perceived Sociability determining Social Presence;
- Perceived Sociability determining Perceived Enjoyment;
- Perceived Enjoyment determining Perceived Ease of Use;
- Perceived Adaptivity determining Perceived Usefulness.

These interrelations all involve new (non UTAUT) constructs and their general presence demonstrates the validity of our expansion.

The influences of the constructs of Social Influence on Intention to Use and of Trust on Intention to Use (dotted lines in Figure 8.1) were not confirmed by the results of any of the experiments. This means Trust and Social Influence do not have a direct influence on acceptance. This does not mean they have no influence: the influence of other factors could be so dominating that the influence of these constructs can simply not be established. For Social Influence this could mean that during or shortly after the initial test of the system either acceptance of this type of technology is not being influenced by any social pressure or that elderly people are not very much influenced by it. When it comes to actual use, however, Social Influence could come into play, as it did in experiment 4.

Regarding the questionnaire: in order to obtain values for Cronbach's Alpha that were high enough, we had to omit items 10 (FC3), 21 (PENJ5), 22 (PEOU1), 37 (SI3) and 45 (TR3) from the results of almost all our experiments. The only exceptions were the following:

- Item 10 (FC3) was not excluded for Experiment 4. The value for Cronbach's Alpha was too low, but would still have remained too low if the item were excluded. Moreover, the concerning construct of Facilitating Condition was not relevant here, because it is hypothesized to determine Use, which was not measured in this experiment.
- Item 21 (PENJ5) was not excluded from the results of Experiment 3, because Cronbach's Alpha was high enough (.846). However, if it would have been excluded, the value for Cronbach's Alpha would have been higher (.874).

We conclude that despite these exceptions, we can delete these items permanently from our questionnaire. The definite list of statements is presented in Appendix F.

8.3 Combined results

The systems used in the experiments were very different and our RoboCare experiment confronted participants not even with an actual system but with movies of a robot. Nevertheless, in experiments 3 to 6, we used the same model and the same questionnaire. In these experiments, the model has shown to explain acceptance of these different systems in different settings, which is consistent with our conjecture in Section 1.5. Therefore we will combine the results of the experiments, thus obtaining a set of data with 188 cases (total number of participants in experiments 3-6).

A regression analysis on these combined results, testing the model validation hypotheses, can give us a further indication on the possibility to generalize the model. As table 8.3 and Figure 8.2 show, many, but not all construct interrelations are confirmed by this analysis (note that for the regression on Use, only the data of experiments 3 and 4 could be used). The R^2 value on Intention to Use is just .59, which is low compared to the scores for the individual experiments (Table 8.1). This is due to the increasing variance caused by combining the results from different data sets.

Hypothesis		*Independent*	*Dependent*	*Beta*	*t*	*Sig*
VH1	*(a)*	*ITU*		*.584*	*4.718***	*.000*
	(b)	*SI*	*† Usage(min)*	*-.096*	*-.803*	*.426*
	(c)	*FC*		*-.126*	*-1.072*	*.288*
		Model: R^2=.29; F=7.600; df=3,184; P=.000				
VH2	*(a)*	*PU*		*,433*	*6,831***	*,000*
	(b)	*PEOU*		*,349*	*2,375**	*,021*
	(c)	*ATT*	*ITU*	*,327*	*4,531***	*,000*
	(d)	*PENJ*		*,134*	*2,070**	*,040*
	(e)	*SI*		*-,120*	*-1,130*	*,264*
	(f)	*Trust*		*,003*	*,042*	*,967*
		Model: R^2=.59; F=37.989; df=6,181; P=.000				
VH3	*(a)*	*ANX*		*,040*	*,565*	*,573*
	(b)	*ATT*	*PU*	*,274*	*3,366***	*,001*
	(c)	*PAD*		*,285*	*3,406***	*,001*
	(d)	*PEOU*		*,403*	*5,364***	*,000*
		Model: R^2=.57; F=31.358; df=4,183; P=.000				
VH4	*(a)*	*ANX*	*PEOU*	*,143*	*2,133**	*,034*
	(b)	*PENJ*		*,429*	*6,379***	*,000*
		Model: R^2=.19; F=13.950; df=2,185; P=.000				
VH5	*(a)*	*PS*	*PENJ*	*,604*	*9,898***	*,000*
	(b)	*SP*		*,147*	*2,401**	*,017*
		Model: R^2=.48; F=82.936; df=2,185; P=.000				
VH6		*Trust*	*PS*	*.394*	*5.842***	*.000*
		Model: R^2=.23; F=34.130; df=1,186; P=.000				
VH7		*PS*	*SP*	*.485*	*7.572***	*.000*
		Model: R^2=.24; F=57.333; df=1,186; P=.000				

Table 8.3. Regression analysis on combined results concerning the model validation hypotheses (section 7.1).
† For Usage only the results of experiments 3 and 4 could be used.

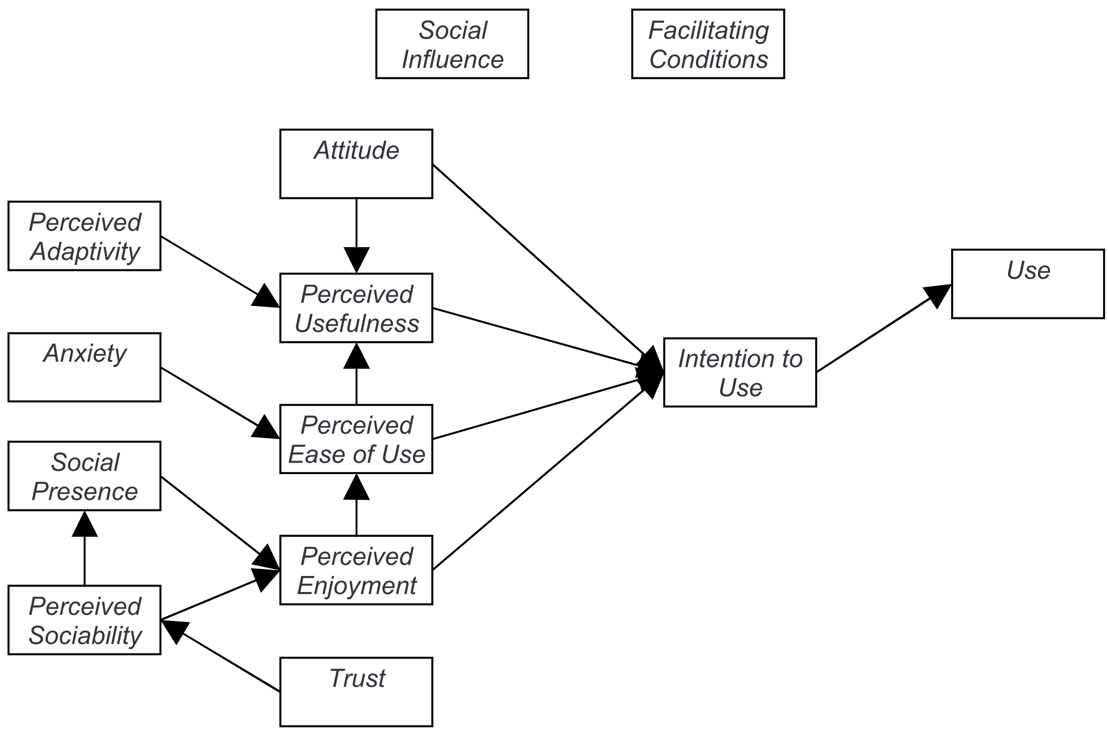


Figure 8.2. Confirmed construct interrelations for combined results

8.4 Path analysis

With a set of 188 cases, we have the possibility of applying structural equation modeling to carry out a path analysis – a type of regression analysis in which the whole model can be tested at once instead of the individual hypotheses. Usually this requires at least 15 cases per independent variable (Ding et al. 1995; Schumacker and Lomax 1996; Gefen et al. 2000) which in our case means we need at least 150 cases (participants).

Testing our model, using the statistical package SPSS AMOS leads to the results pictured in Figure 8.3. It shows that most of the construct interrelations have significant weights (a value of .08 or higher is significant at the 5% level), but there are two exceptions. First, Anxiety determining Perceived Usefulness is not strong enough (and Anxiety determining Perceived Ease of Use is very weak). Second, Trust determining Intention to Use is not strong enough – note that this hypothesized influence has not been confirmed by any regression analysis.

Figure 8.3 also presents the Chi-square, the degrees of freedom (which is 66 sample moments minus 26 estimated parameters) and the goodness-of-fit index (GFI), which indicates how well the dataset fits the model. This GFI of .79 is actually quite low, it is recommended to have it close to .90 (Tanaka 1987) and chi-square divided by degree of freedom ($\chi 2/df$) should be below 5 (Hayduk 1987), which is also not the case. This made us explore other not yet established construct interrelations that would be plausible and would raise the GFI.

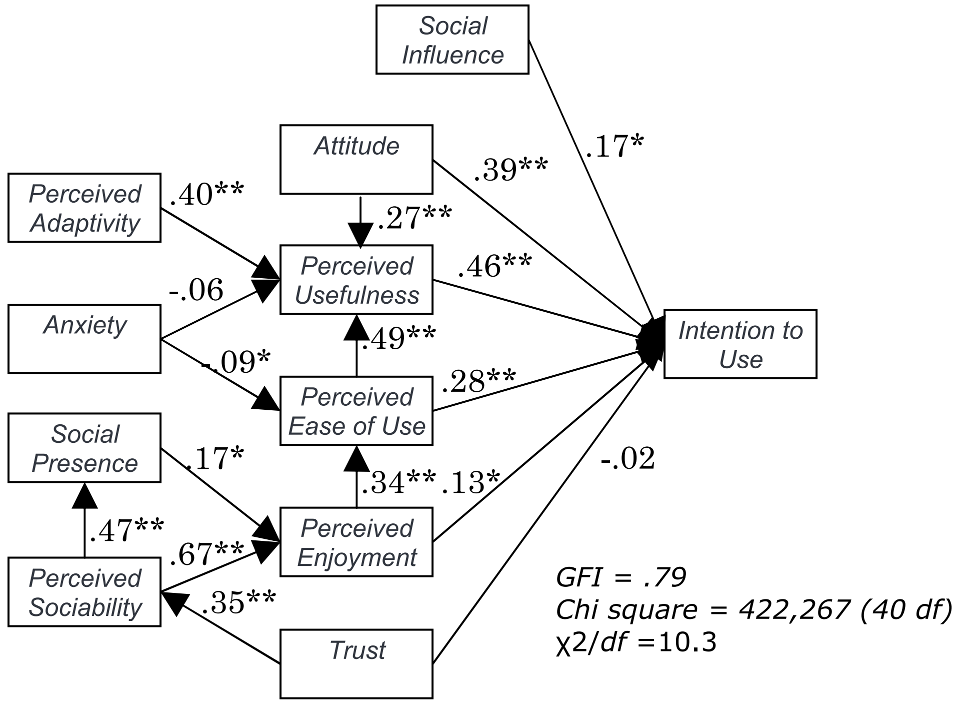


Figure 8.3. Model path estimates

A common way to do such an exploratory path analysis, is to calculate modification indices (Chou and Bentler 1990; Mitchell 1992; Schumacker and Lomax 1996; Schubert et al. 1999; Scheiner et al. 2000), which can also be done with SPSS AMOS. A value in these modification indices above 3.84 suggests that adding that path may significantly improve model fit (Hair et al. 1998), since 3.84 is the critical value of the chi-square statistic with 1 degree of freedom at the 5% significance level.

Table 8.4 shows modification indices we found to fit this criterion for interrelations that were plausible: Social Influence, Perceived Adaptivity and Anxiety determining Attitude and for Attitude determining Trust. When applying these suggested relations to our model we found highly significant weight scores ($P < .005$).

.

Construct interrelations			**Modification indices**
ATT	<--	SI	48.325
ATT	<--	ANX	6.183
ATT	<--	PAD	28.612
Trust	<--	ATT	48.830

Table 8.4. Modification indices.

Figure 8.4 shows the model diagram including the additional interrelations. It now has a GFI of .96, which is high (the value for degrees of freedom is now 66 sample moments minus 29 estimated parameters) and the value for χ^2/df is now clearly below 5.

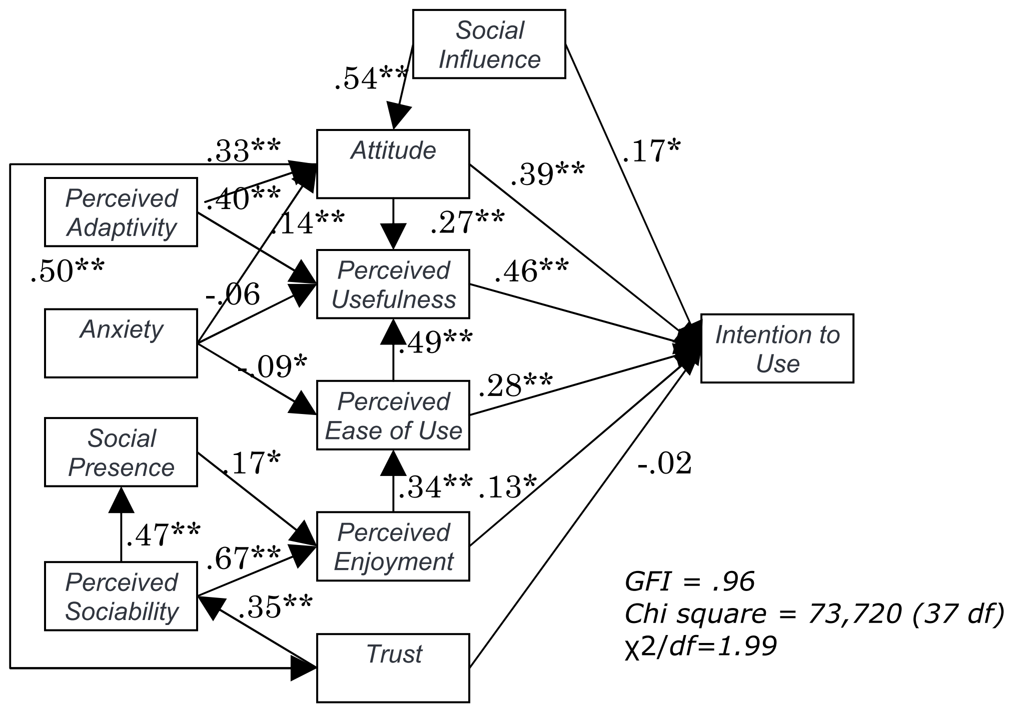


Figure 8.4. Model path estimates with additional interrelations

We carried out an additional linear regression analysis for these new interrelations. The results are presented in Table 8.5.

Independent	Dependent	Beta	t	Sig
PAD		.249	3.927**	.000
SI	ATT	.423	6.682**	.000
ANX		.178	2.984**	.003
ATT	Trust	.511	8.108**	.000

Table 8.5. Alternative construct interrelations for combined results

As with other construct interrelations, the impact of the three constructs on Attitude differs for each system, setting and condition (see Table 8.7) as does the impact of Attitude on Trust (Table 8.6).

Experiment/System	Beta	t	Sig
3/iCat touch screen	.622	4.354**	.000
4/Steffie - screen agent at home	.624	4.229**	.000
5a/iCat WOZ - More social	.718	4.371**	.000
5b/iCat WOZ - Less social	.574	3.056*	.006
6a/Robocare movie, not adaptive, not adaptable	.650	3.820**	.001
6b/Robocare movie, adaptable	-.067	-.278	.784
6c/Robocare movie, adaptive, user controlled	.618	3.511**	.002
6d/Robocare movie, adaptive, not user controlled	.446	2.231*	.037

Table 8.6. Regression analysis of Attitude determining Trust for different systems

Experiment/System	***Independent***	***Beta***	***t***	***Sig***
3/iCat touch screen	*PAD*	*.303*	*2.408**	*.023*
	SI	*.212*	*1.691*	*.102*
	ANX	*.553*	*4.353***	*.000*
4/Steffie	*PAD*	*.506*	*4.443***	*.000*
screen agent at home	*SI*	*.417*	*3.567***	*.001*
	ANX	*.242*	*2.372**	*.025*
5a/iCat WOZ	*PAD*	*.506*	*2.267**	*.038*
More social	*SI*	*.248*	*1.093*	*.290*
	ANX	*-.347*	*-2.093*	*.053*
5b/iCat WOZ	*PAD*	*.448*	*2.461**	*.025*
Less social	*SI*	*.261*	*1.468*	*.160*
	ANX	*-.278*	*-1.572*	*.134*
6a/Robocare movie	*PAD*	*.311*	*1.620*	*.123*
Not adaptive.	*SI*	*.445*	*2.319**	*.032*
not adaptable	*ANX*	*.251*	*1.305*	*.208*
6b/Robocare movie	*PAD*	*-.135*	*-.516*	*.613*
Adaptable	*SI*	*.585*	*2.372**	*.031*
	ANX	*-.091*	*-.388*	*.704*
6c/Robocare movie	*PAD*	*.066*	*.370*	*.716*
Adaptive	*SI*	*.682*	*3.693***	*.002*
user controlled	*ANX*	*.135*	*.727*	*.477*
6d/Robocare movie	*PAD*	*-.148*	*-.750*	*.463*
Adaptive	*SI*	*.559*	*2.841**	*.011*
not user controlled	*ANX*	*.062*	*.319*	*.754*

Table 8.6. Regression with new determinants of Attitude for different systems

All hypothetical construct interrelations have been confirmed at some point, either by a linear regression analysis or by path analysis, except for the impact of Trust on Intention to Use, which has never been confirmed. This is why we decided to leave this interrelation out of the final model.

8.5 Conclusions

In this chapter we presented an overview of the results of our experiments and concluded that we have developed a model with additional constructs with a higher explanatory power than the originally used UTAUT model. A path analysis on the combined results of our experiments with the new model confirmed our model, but we had to refine our model to establish an acceptable goodness of fit. We learned that Trust is moderated by Attitude and that Attitude is moderated by Social Influence, Perceived Adaptivity and Anxiety.

The model that emerges from this is visualized in Figure 8.5. We called it ‘The Almere model’: since most of the experiments have been carried out in the Dutch city of Almere (either in eldercare institutions or at the homes of elders still living independently), we decided to honor our participants and collaborating staff by naming the model after this city.

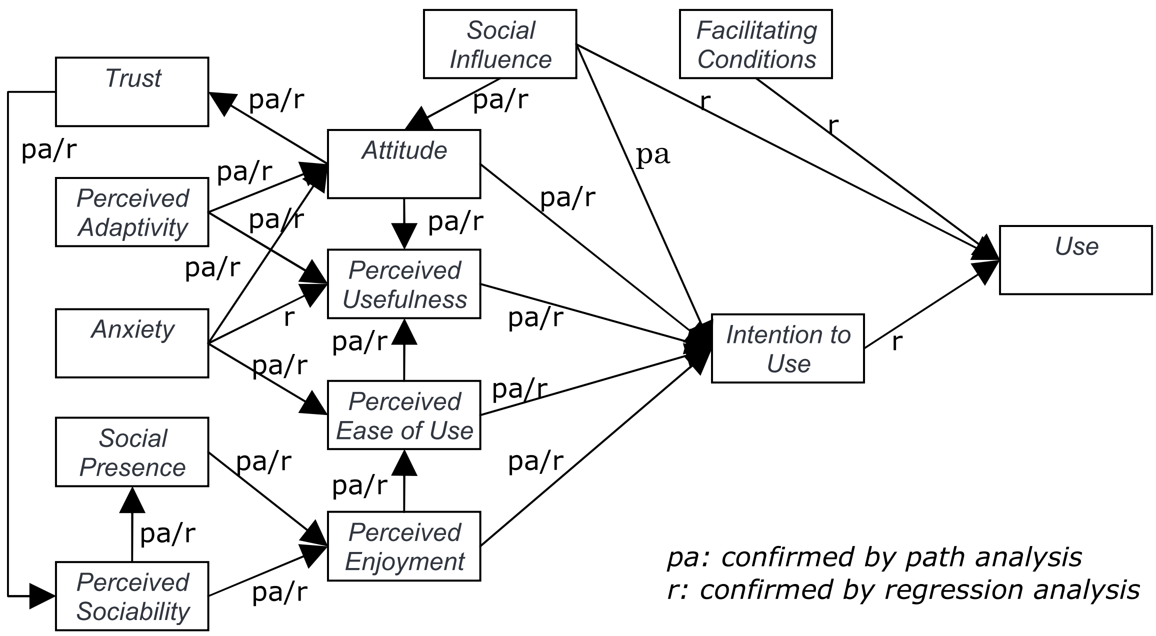


Figure 8.5. Confirmed construct interrelations after regression and exploratory path analysis

We found that this model reflects system differences by showing differences in confirmed construct interrelations. This concerns in particular the influences on Intention to Use: the more utilitarian systems show a predominating influence of Perceived Usefulness and the more hedonic systems show a more predominating influence of Perceived Ease of Use and Perceived Enjoyment.

The model differs from the one we developed solely based on the hypotheses that were confirmed by regression analyses. First, we added construct interrelations concerning the determination of Attitude and Trust. Second, there are two differences concerning the confirmation of originally (in section 4.5) hypothesized interrelations:

1. The influence of Anxiety on Perceived Usefulness is not strong enough to be significant in a path analysis. This confirms our regression analysis with the combined scores (see Table 8.2 and Figure 8.2). However, since we confirmed this interrelation in both experiment 4 (Steffie) and 5 (iCat in two conditions), concerning two very different systems and settings, we will keep it in our model.
2. Social Influence is a significant moderator of Intention to Use in our path analysis. While this interrelation was never significant in any of the regression analyses (though often not far from being significant), in our path analysis it is no longer suppressed by other influences.

Comparing our new model to the original UTAUT model, we find that we added five constructs:

- Perceived Enjoyment, a construct related to the hedonic nature of the used technology.

- Perceived Adaptivity, a construct related to the specific user group we are addressing.
- Trust, Social Abilities and Social Presence, three constructs related to technology that can be perceived as a social entity.

9. Summarized findings and final conclusions

9.1 Introduction

In this final chapter we close this thesis by summarizing our findings, reflecting on the research questions and pointing out our contributions. We will conclude with a discussion section that contains suggestions for further research.

9.2 Reflection on the research questions

Our main research question drove us to find a method to explain and predict the influences on acceptance of assistive social robots by elderly users. We therefore answered the following sub-questions:

1. To what extent is the most prevailing technology acceptance model able to explain and predict acceptance of assistive social robots by elderly users?
2. If the most prevailing model is not able to adequately explain and predict acceptance of assistive social robots by elderly users, can we set up a new model by incorporating new influences and prove it to have a better explanatory power than that model?
3. What evidence can be found concerning alternative influences on acceptance of assistive social robots by elderly users?

The first question has been addressed by testing the UTAUT model on two different systems, each in a more and less social condition (see Chapter 3). We found that the model did not perform very well: on average it explained only 47 percent of the variance in Intention to Use. Moreover, for the more social conditions only 34% percent of the variance in Intention to Use was explained. We concluded that the model needed further adaptations, expecially concerning the more social aspects of acceptance.

We thus addressed the second and third question by adding new constructs based on related literature, our own observations and (finally) exploratory path analysis. This resulted in our ‘Almere model (presented in Chapter 4)’, which indeed performed better: it explained on average 71 percent of variance in the Intention to Use and a path analysis resulted in a goodness of fit of .96 (see Capter 8). Nevertheless, it has some shortcomings that need to be addressed in further research (see next session).

If we recall the demands that we set, we may conclude that our model suffices:

1. It should have the ability to explain acceptance under a wide variety of experimental conditions.
2. It should show robustness in quantitative results.
3. It should aim to identify all influences on acceptance of this type of technology for this user group.

The model has been used in different conditions, systems and settings used for experiments 3 to 6 and in all these experiments it performed well, explaining between 62 and 88 percent of variance in the Intention to Use. The robustness was demonstrated by the high Cronbach's Alpha scores.

The aim to identify all influences cannot be fully established, since the explained variance in the Intention to Use is not 100 percent. However, the above mentioned percentages indicate we have identified direct and indirect influences sufficiently to explain acceptance of this type of technology for this user group

We can therefore state that the Almere model has predictive strength and constructs that have proven to be reliable in various settings, and we have seen a satisfying goodness of fit score on our data. We have thus developed and validated a model that can be used to predict and explain acceptance of assistive social robots. If applied to different systems and settings it results in varying dominating influences. These varying influences can in fact explain the impact of these system differences. The fact that the systems we used had varying functionalities confirms findings of Van der Heijden (2004) and Chesney (2006) that that 'purpose of use' is essential in determining the factors that predict acceptance.

9.3 Main contributions

We have developed and validated a model (Figure 9.1) that can be used to predict and explain acceptance of assistive social robots. If applied to different systems and settings it results in a difference in dominating influences that can be used to explain the impact of these differences. Moreover, it can be used to establish differences in the perception of aspects of this technology between user groups or individuals.

In the process of building this model, we were able to demonstrate that, when dealing with robot acceptance, we are not only dealing with acceptance of technology, but also acceptance of what is perceived as a social entity. We therewith addressed social acceptance of robots as opposed to functional acceptance. Moreover, we introduced conversational acceptance as an aspect of social acceptance that can be measured not only with a questionnaire, but also by observing user behavior.

By demonstrating that social robots and screen agents are perceived as social entities, we established that elderly users may view a robot as a person and may even be open to being touched by this entity. Even when they are fully aware that they are talking to something that is just a construction of metal and plastic, they behave like they are talking with something more than that. And in addition: the more sociable this system is, the more they experience it to be a social entity and the more they enjoy it.

Very close to social abilities is adaptivity: we were able to demonstrate how older adults expect a robot to be adaptive to their needs. They want a system to be helpful with something they need – and not with something they do not need yet. But when being adaptive, the robot should still be under their control.

The model introduces some constructs that are very new in the field of acceptance technology: never before were Social Abilities, Perceived Adaptivity and Social Presence part of a technology acceptance model. In this thesis we established their relevance in the context of social assistive agents used by older adults: they actually make an acceptance model more effective. These constructs can possibly also be applied in a broader context in future research considering either assistive technology or social robots in general.

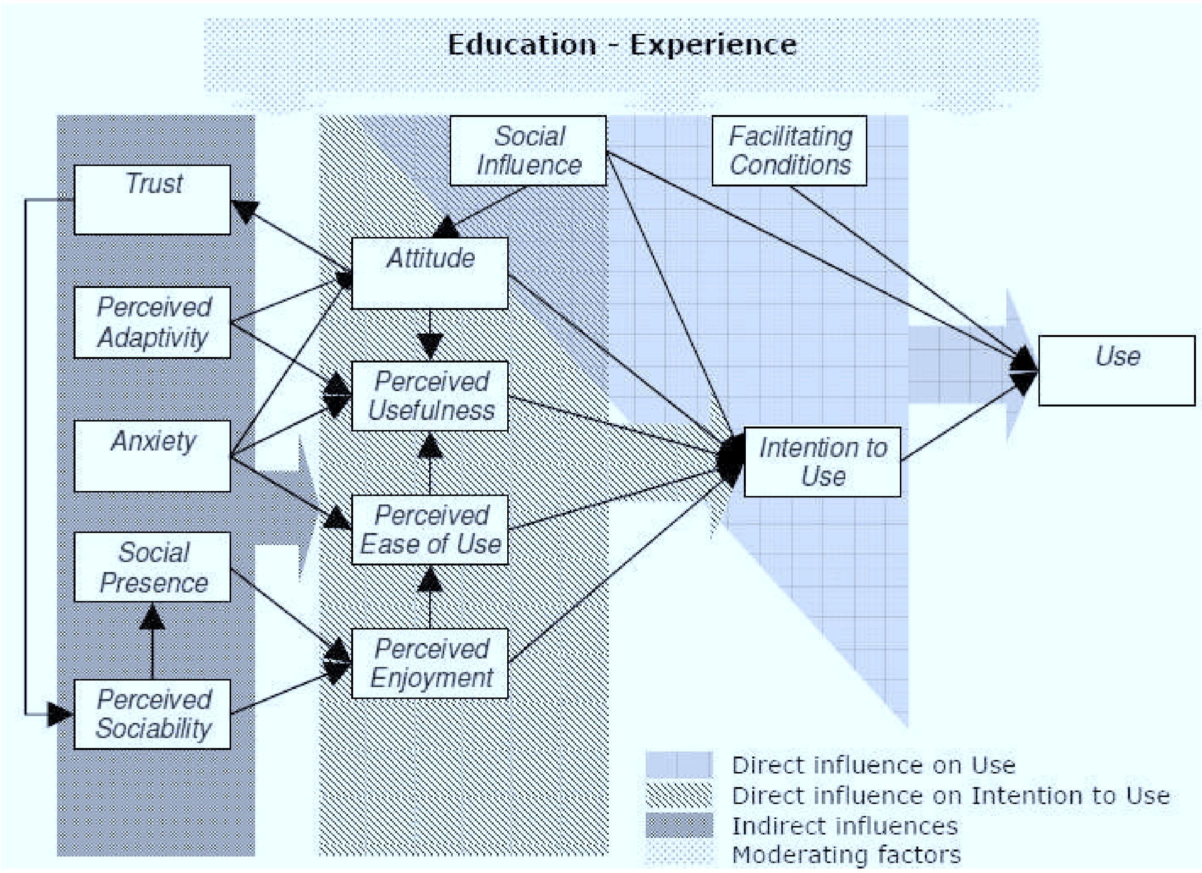


Figure 9.1. Developed model

The regression results of the four experiments with this model demonstrated how the model reflects the differences between the systems, settings and tasks:

- A system that had more social and entertainment features (e.g. the iCat in Experiment 3), shows a stronger influence of Perceived Enjoyment and Perceived Ease of Use.
- A system that has more assistive functionalities (RoboCare, Steffie), shows a stronger influence of Perceived Usefulness.

The applicability of the model as a refined explanatory instrument benefits from the *indirect* influences that are incorporated. As the UTAUT model only features direct influences on Intention to Use and actual use, it lacks the possibilities to explain the differences in strength of these influences. For example, a stronger influence of Perceived Usefulness may remain unexplained in UTAUT, while it could be explained by a strong influence of Perceived Adaptivity, Anxiety, Attitude or Perceived Ease of Use – and this could concern just one or a combination of these constructs.

9.4 Discussion and further research

We tested our model on systems that differed in many aspects: zoomorphic versus anthropomorphic, screen agent versus robot, mobile versus non-mobile - and besides these differences on embodiment there were also different tasks and settings. As stated earlier, this enlarges the possibility to generalize the model and it enabled us to examine how the model would reflect these differences. However, a study that would explicitly focus on system or task differences to compare participant scores (e.g. to examine preferences) would demand systems that just differ in one embodiment or task aspect at a time. This would enable us to research specific acceptance questions on the influence of embodiment, on adding functionalities, on the environment in which the system is used, or changing (social) behavior of the system. Specifically for social robots, this would make it possible to study the effect of adding or removing specific social abilities: we found that they are generally relevant when studying acceptance, but we were not able to specify the contribution of specific abilities (e.g. gaze, expressiveness, apologizing).

Also the different constructs could be subject to further research. For example, our model development process revealed the importance of Attitude in this context. It was one of the most significant influences on Intention to Use in three of the four experiments and it appeared to have additional interrelations after our explanatory analysis in Chapter 8. With this finding we are consistent with that of Yang and Yoo [31], that Attitude is a crucial factor that needs further investigation and perhaps specification. Their idea that there are different types of attitude (cognitive and affective) could actually be applied very well to robot technology, because of its dual nature, combining technical and social entities.

Furthermore, we included some moderating factors, but since they were not in the focus of our interest (they do not directly contribute to the explanatory power of a model), there remains much to be researched on this specific technology and for this specific user group. We only included the moderating factor hypotheses that were presented in the UTAUT model and that were applicable to our context, while there is much research on alternative influences of the incorporated factors and on alternative factors (for an overview, see Sun and Zhang 2006). There is, for example, research that indicates voluntariness to be a generally relevant factor (Schaper and Pervan 2007; Wu et al. 2007). Since in our experiments all participants were people that volunteered for it, we had no data to confirm this. Of course, voluntariness can be expected to be more of influence in a working environment, where people can be obliged to use a technology. However, also in an eldercare environment motivation can be a factor and if this is found to be true, additional research should establish if acceptance is determined similarly by older adults who are less willing to try this new technology.

Another item for further research concerns our behavioral analysis. As we stated in Chapter 5, we think the question should be addressed whether specific conversational expressions occurred in response to similar expressions by the robot (a smile in response to a smile, a frown in response to a frown). In that case we would be speaking of imitative behavior. This would be the occurrence of a well known phenomenon in psychology called the chameleon effect (Chartrand and Bargh 1999). It concerns imitative behavior between humans, which seems to occur naturally unless two people do not like each other. The occurrence of this behavior could even very well be interpreted as a sign of acceptance (Kahn et al. 2006). But during behavior analysis the observers just counted the number of behaviors, without looking at the behavior of the robot that evoked it - the camera was always directed towards the participant. In future research this possibility of imitative behavior could be something to observe, also when comparing agents with different embodiments, since it could add interesting viewpoints to HRI theory on this aspect (Dautenhahn and Nehaniv 2002).

Also our study would benefit from an experiment that would include repetitive testing. This would make it possible to study factors that are often subject to change within individuals as they become more experienced with or more informed about technology, like it has been established for Attitude (Macer and Ng 2000; Morris and Venkatesh 2000; Bhattacherjee and Premkumar 2004) and Trust(Grandison and Sloman 2000; Lippert and Davis 2006). Moreover, these two influences may change when aging adults develop different mental and physical needs. This would, according to our model, lead to a change in Intention to Use: a hypothetical effect that demands to be confirmed by a long term study.

In addition, our claims of validity would benefit from a study which includes a path analysis on data concerning one specific system. We realize however, that this demands a high quantity of participants, which is hard to realize with this type of technology and user group.

Furthermore, although we did use different systems, our study would benefit from more studies with different assistive social robots that have features that our test robots lacked. The iCat is for example not mobile and nor iCat, nor RoboCare are capable of carrying objects or people.

Moreover, despite the excellent goodness of fit and predictive strength, in some aspects our model still needs further research before we can claim its completion. First of all, usage is just measured over a short period in experiments 5 and 6. Studies involving a longer usage period (rather months than days or weeks) can give more insight on the actual use of this technology.

Finally, the scope of this study has been assistive social robots and screen agents. However, as we pointed out in our introduction, the development of this technology is part of a broader field of assistive technology that enables aging adults to live more independently. Positioning our study within this broader perspective initiates two ideas. First, many of the findings can possibly be generalized: the influence of adaptivity on acceptance, the contribution of experience, the relevance of enjoyment and even the experience of sensing a social entity in a socially interactive system. Secondly, we may realize that robots and screen agents can have a specific intermediating function, being the social actors amidst the advanced technologal environment – for which they will often be the interface. Making these social actors as acceptable as we can, will inevitably be a crucial step in making the surrounding assistive technology acceptable.

Samenvatting (*Dutch summary*)

In dit proefschrift staat de vraag centraal wat de factoren zijn die de acceptatie bepalen van dienstverlenende, sociale robots door ouderen (65+), opdat we deze acceptatie kunnen voorspellen en verklaren. Deze robots kunnen een belangrijke rol gaan spelen binnen intelligente omgevingen waarin ouderen langer op zichzelf blijven wonen. Daarmee is dit een ontwikkeling die bij kan dragen aan de aanpak van de vergrijzingsproblematiek die in de komende decennia steeds meer druk op de Westerse samenleving zal uitoefenen. Cruciaal daarbij is dat die robots daarbij volledig geaccepteerd worden door de gebruikers. Juist bij deze specifieke doelgroep is dit een precaire kwestie: er zijn voorbeelden van technologie die in theorie veel voor ouderen kan betekenen, maar die vervolgens niet, of niet volledig geaccepteerd wordt.

Om deze acceptatie meetbaar, voorspelbaar en verklaarbaar te maken, gebruiken we een methodiek die *technology acceptance modeling* (TAM) heet en die tot nu toe vooral gebruikt wordt bij het voorspellen en verklaren van acceptatie van ICT in een werkomgeving. Deze methodiek houdt in dat men in kaart brengt wat de invloeden zijn op acceptatie van een bepaalde technologie door een bepaalde gebruikersgroep. Die invloeden zijn vooral de aspecten van de technologie als gebruiksgemak en bruikbaarheid zoals die ervaren worden door de gebruikers in een test. De onderzoekers proberen deze invloeden meetbaar te maken met als belangrijkste instrument een lijst met vragen. Deze lijst is zodanig samengesteld dat er per invloed minimaal twee vragen zijn, die samen een zogenaamd construct vormen. De resultaten hiervan kunnen kwantitatief worden verwerkt met verschillende statistische technieken (vooral correlatie, regressie en *structural equation modeling*).

Echter, deze TAM-methodiek is niet zonder meer toepasbaar als het gaat om sociale robots en ouderen. We laten dat in dit proefschrift zien door twee experimenten te beschrijven waarin we een TAM-methodiek gebruiken die pretendeert alle relevante invloeden in kaart te brengen: UTAUT - Unified Theory of Acceptance and Use Technology. We passen dit toe op een sociale robot en een sociale *screen agent* (een geanimeerd personage op een beeldscherm). We manipuleren daarbij de sociale vaardigheden bij beide systemen, zodat er een meer en minder sociale variant van elk systeem is.

Het model blijkt in deze context weinig voorspelbaar vermogen te hebben en niet bruikbaar te zijn voor het verklaren van verschillen tussen de beleving van ouderen bij een meer en minder sociale robot. Dit is opmerkelijk omdat uit aansluitende metingen (gedragsobservaties en vragen buiten het model) blijkt dat die verschillen wel degelijk van invloed zijn op gedrag en perceptie van de gebruikers.

We zoeken de verklaring in het feit dat deze methodiek zich traditioneel vooral richt op 'functionele' acceptatie. We stellen dat een model waarin ook gebruiksplezier en 'sociale' acceptatie opgenomen zijn, wellicht beter voldoet in deze context: het voorspelt beter en kan effectiever gebruikt worden om verschillen tussen systemen en/of gebruikers te verklaren. We ontwikkelen vanuit deze gedachte een model dat naast traditionele TAM invloeden als bruikbaarheid, bedieningsgemak en attitude ook niet utilitaire (plezier bij het gebruiken) en vooral sociale aspecten van de technologie bevat (zoals sociale vaardigheden, vertrouwen en het gevoel een sociale entiteit te ontmoeten). Bovendien nemen we – op basis van bevindingen in verwant onderzoek - adaptiviteit op, als een invloed die specifiek is voor technologie voor ouderen.

We rechtvaardigen vervolgens het opnemen van deze nieuwe invloeden met specifiek hierop gerichte experimenten. Allereerst zetten we wederom een experiment op waarbij er een meer en minder sociale versie van dezelfde robot gebruikt wordt voor twee gescheiden groepen gebruikers. We tonen hierbij aan dat de sociale versie inderdaad als socialer ervaren wordt en dat de acceptatie – zij het indirect – positief wordt beïnvloed door de geïmplementeerde sociale vaardigheden. Bovendien voorspelt het nieuwe model de intentie om het systeem te gaan gebruiken aanmerkelijk beter.

Vervolgens voeren we een experiment uit waarin adaptiviteit gemanipuleerd wordt. We maken daarbij gebruik van video's van een oudere gebruiker met een sociale robot. Deze video's worden getoond aan verschillende groepen gebruikers in vier varianten: neutraal, aanpasbaar, adaptief (zichzelf aanpassend) met gebruikerscontrole en adaptief zonder gebruikerscontrole. Vervolgens gebruiken we het model om de respons van gebruikers te verzamelen en verschillen te analyseren. De hoogste scores zijn voor de versie die adaptief is maar met gebruikerscontrole (middels een vraag om goedkeuring). Dit experiment laat zien dat adaptiviteit inderdaad van invloed is op de acceptatie en dat ook voor deze setting het model de intentie tot gebruik bevredigend voorspelt.

Hierna valideren we het gehele model door analyse van de resultaten van experimenten waarbij niet alleen de intentie tot gebruik, maar ook het daadwerkelijke gebruik over een periode van zeven tot tien dagen wordt gemeten. We doen dit met een robot die aangestuurd kan worden door een aanraakscherm en een screen agent die bij gebruikers thuis op de PC wordt geïnstalleerd. In beide gevallen blijkt de uitgesproken intentie tot gebruik het daadwerkelijke gebruik te voorspellen. Dit geeft aan dat we een bruikbaar instrument hebben.

Ten slotte inventariseren we de uitkomsten van alle verschillende experimenten en stellen definitief vast welke invloeden direct de intentie tot gebruik beïnvloeden en welke invloeden indirect zijn. Hoewel de voorspellende kracht van het uiteindelijke model verschilt per systeem, voorspelt het beter en kan effectiever gebruikt worden om de invloed op de perceptie en acceptatie van verschillen tussen systemen te verklaren dan het initieel gebruikte UTAUT-model.

Met dit model hebben we aldus de invloeden op acceptatie van sociale robots door ouderen in kaart gebracht. We hebben daarbij laten zien dat ook sociale processen hierbij een rol spelen: de gebruikers ervaren niet alleen de technologie, maar ook een sociale entiteit. Om dit in het model te kunnen opnemen, hebben we een model gebouwd met invloeden die nooit eerder in een acceptatiemodel zijn voorgekomen.

We hebben hiermee een instrument ontwikkeld waarmee de perceptie van een robot of screen agent kwantitatief kan worden onderzocht. Hiermee kunnen verschillen tussen systemen en omstandigheden en tussen individuele gebruikers of gebruikersgroepen in kaart worden gebracht. Het model kan met die mogelijkheden gebruikt worden in de verschillende fases van het ontwikkelproces van een sociale robot of screen agent. Bovendien kan het worden toegepast bij de implementatie van sociale robots en screen agents in de technologische gebruiksomgeving.

Met die gebruiksomgeving zijn we ook weer dicht bij ons uitgangspunt. De bevindingen van deze studie hebben immers betrekking op robots en screen agents voor ouderen in een ondersteunende omgeving waarin ze langer zelfstandig kunnen leven. Veel van de inzichten die voortkomen uit ons onderzoek kunnen dan ook breder toegepast worden: de invloed van adaptiviteit op acceptatie, het belang van gebruiksplezier en het ervaren van een sociale entiteit bij de interactie met een systeem. Bovendien ligt de rol van sociale robots en screen agents vooral binnen deze omgeving, waar ze met hun mogelijkheid tot sociale interactie een schakel kunnen vormen tussen technologie en gebruiker. Dit impliceert ook, dat we met het werken aan de acceptatie van sociale robots in feite werken aan de acceptatie van een toekomstige ondersteunende technologische omgeving voor ouderen.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes* 50(2): 179-211.

Allwood, J. and L. Cerrato (2003). A study of gestural feedback expressions. *Proceedings of the First Nordic Symposium on Multimodal Communication*, Copenhagen: 7-20

Alpert, S. R., J. Karat, C. M. Karat, C. Brodie and J. G. Vergo (2003). User attitudes regarding a user-adaptive eCommerce web site. *User Modeling and User-Adapted Interaction* 13(4): 373-396.

Axelrod, L. and K. Hone (2005). Identifying Affectemes: Transcribing Conversational Behaviour. *Proceedings of the Symposium on Conversational Informatics for Supporting Social Intelligence and Interaction*, Salt Lake City, Utah, USA : 430-439

Axelrod, L. and K. Hone (2006). Affectemes and allaffects: a novel approach to coding user emotional expression during interactive experiences. *Behaviour & Information Technology* 25(2): 159 - 173

Bahadori, S., A. Cesta, G. Grisetti, L. Iocchi, R. Leone, D. Nardi, A. Oddi, F. Pecora and R. Rasconi (2003). RoboCare: an Integrated Robotic System for the Domestic Care of the Elderly. *Proceedings of workshop on Ambient Intelligence AI* IA-03, Pisa, Italy*.

Bahadori, S., A. Cesta, G. Grisetti, L. Iocchi, R. Leone, D. Nardi, A. Oddi, F. Pecora, R. Rasconi and C. N. delle Ricerche (2003). RoboCare: Pervasive Intelligence for the Domestic Care of the Elderly. *Intelligenza Artificiale* 1(1): 16-21.

Bailenson, J. N., J. Blascovich, A. C. Beall and J. M. Loomis (2001). Equilibrium revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators and Virtual Environments* 10: 583-598.

Baltus, G., D. Fox, F. Gemperle, J. Goetz, T. Hirsch, D. Magaritis, M. Montemerlo, J. Pineau, N. Roy and J. Schulte (2000). Towards personal service robots for the elderly. *Proceedings of the Workshop on Interactive Robotics and Entertainment (WIRE-2000)*.

Banerjee, M., M. Capozzoli, L. McSweeney and D. Sinha (1999). Beyond kappa: A review of interrater agreement measures. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*: 3-23.

Barea, R., L. M. Bergasa, E. López, M. S. Escudero, J. A. Hernández and Y. Willemaers (2004). Tele-medicine system based on a personal robotic assistant. *Proceedings 10th IEEE International Conference on Methods and Models in Automation and Robotics*, Miedzyzdroje, Poland: 909-915

Bartlett, B., V. Estivill-Castro, S. Seymon and A. Tourky (2003). Robots for preorientation and interaction of toddlers and preschoolers who are blind. *Proceedings from the Annual Conference of the American Cultural Resources Association*, Dallas, USA: 338-346

Bartneck, C., J. Reichenbach and A. J. N. van Breemen (2004). In your face, robot! The influence of a character's embodiment on how users perceive its emotional expressions, Design and Emotion. *Proceedings of the Design and Emotion 2004 Conference*, Ankara, Turkey: 77-82

Bartneck, C., T. Suzuki, T. Kanda and T. Nomura (2007). The influence of people's culture and prior experiences with Aibo on their attitude towards robots. *AI & Society* 21(1-2): 217-230.

Beck, A., Edwards, N., Friedman, B., Khan, P. (2003). Robotic Pets and the Elderly. from http://www.ischool.washington.edu/robotpets/elderly/.

Beck, A. and A. Katcher (1996). *Between Pets and People*. West Lafayette, Purdue University Press.

Benbasat, I. and H. Barki (2007). Quo vadis, TAM. *Journal of the Association for Information Systems* 8(4): 211-218.

Benyon, D. (1993). Adaptive systems: a solution to usability problems. *User Modeling and User-Adapted Interaction* 3(1): 65-87.

Bhattacherjee, A. and G. Premkumar (2004). Understanding changes in belief and attitude toward information technology usage: A theoretical model and longitudinal test. *MIS Quarterly* 28(2): 229-254.

Bickmore, T., Caruso, L., and Clough-Gorr, K. (2005). Acceptance and Usability of a Relational Agent Interface by Urban Older Adults. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, Portland, USA: 1212-1215

Bickmore, T., Picard R. W. (2004). Towards Caring Machines. *Proceedings of CHI '04*, Vienna: 1489-1492

Bickmore, T. and D. Schulman (2006). The Comforting Presence of Relational Agents. *Proceedings of CHI '06*, Montréal, Québec, Canada: 550-555

Bickmore, T. and D. Schulman (2007). Practical approaches to comforting users with relational agents, Proceedings of *CHI '07*, San Jose, CA, USA: 2291-2296

Bickmore, T. W. (2003). *Relational Agents: Effecting Change through Human-Computer Relationships*, Massachusetts Institute of Technology. Ph.D. Thesis

Bickmore, T. W., L. Caruso, K. Clough-Gorr and T. Heeren (2005). 'It's just like you talk to a friend' relational agents for older adults. *Interacting with Computers* 17(6): 711-735.

Bickmore, T. W., L. Pfeifer, D. Schulman, S. Perera, C. Senanayake and I. Nazmi (2008). Public displays of affect: deploying relational agents in public spaces. *Proceedings of CHI '08*, Florence, Italy: 3297-3302

Biocca, F., C. Harms and J. K. Burgoon (2003). Toward a more robust theory and measure of social presence: review and suggested criteria. *Presence: Teleoperators and Virtual Environments* 12(5): 456-480

Breazeal, C. (2000). *Sociable machines: expressive social exchange between humans and robots*. Department of Electrical Engineering and Computer Science, Boston, MA, MIT.

Breazeal, C. (2002). Regulation and entrainment in human-robot interaction. *International Journal of Robotics Research* 21(10-11): 883-902.

Breazeal, C. (2003). Towards sociable robots. *Robotics and Autonomous Systems* 42(3-4): 167-175

Breazeal, C. and L. Aryananda (2002). Recognition of affective communicative intent in robot-directed speech. *Autonomous Robots* 12(1): 83-104.

Breazeal, C., Brooks A., Gray, J., Hoffman, G., Kidd C., Lee, H., Lieberman, J., Lockerd, A. and Mulanda, D. (2003). Humanoid Robots as Cooperative Partners for People. *International Journal of Humanoid Robots* 2(1): 1-34.

Breazeal, C. L. (2000). *Sociable machines: expressive social exchange between humans and robots*. Dept. of Electrical Engineering and Computer Science. Cambridge, Massachusetts Institute of Technology. PhD thesis.

Broekens, J., M. Heerink and H. Rosendal (2009). The effectiveness of assistive social robots in elderly care: a review. *Gerontechnology journal* 8(2): 94-103.

Burgard, W., A. B. Cremers, D. Fox, D. Haehnel, G. Lakemeyer, D. Schulz, W. Steiner and S. Thrun (1998). *The Interactive Museum Tour-Guide Robot*. Hoboken, NJ, USA, John Wiley & Sons LTD.

Burton-Jones, A. and G. S. Hubona (2005). Individual differences and usage behavior: revisiting a technology acceptance model assumption. *ACM SIGMIS Database* 36(2): 58-77.

Cacioppo, J. T., B. N. Uchino, S. L. Crites, M. A. Snydersmith, G. Smith, G. G. Berntson and P. J. Lang (1992). Relationship Between Facial Expressiveness and Sympathetic Activation in Emotion: A Critical Review, With Emphasis on Modeling Underlying Mechanisms and Individual Differences. *Journal of Personality and Social Psychology* 62(1):110-128

Camarinha-Matos, L. M. and H. Afsarmanesh (2002). Design of a virtual community infrastructure for elderly care. *Proceedings 3rd IFIP Working Conference on Infrastructures for Virtual Enterprises*, Sesimbra, Portugal: 439-452

Cappella, J. N. (1983). Conversational involvement: Approaching and avoiding others. *Nonverbal interaction*. I. J. W. R. H. (Eds.). Beverly Hills, CA: Sage**:** 113-148.
Cappella, J. N. and C. Pelachaud (2001). Rules for Responsive Robots: Using Human Interactions to Build Virtual Interactions. *Stability and Change in Relationships*. H. T. R. A.L. Vangelisti, and M. A. Fitzpatric. New York, Cambridge University Press.
Cerrato, L. A. (2002). Comparison between Feedback Strategies in Human-to-Human and Human-Machine Communication. *Proceedings of ICSLP*, Denver, Colorado, USA: 557-560
Cesta, A., G. Cortellessa, M. Giuliani, F. Pecora, R. Rasconi, M. Scopelliti and L. Tiberio (2007). Proactive assistive technology: An empirical study. *Lecture notes in computer science* 4662: 255-268.
Cesta, A., G. Cortellessa, M. V. Giuliani, F. Pecora, M. Scopelliti and L. Tiberio (2007). Psychological Implications of Domestic Assistive Technology for the Elderly. *PsychNology Journal* 5(3): 229-252.
Cesta, A., G. Cortellessa, F. Pecora and R. Rasconi (2007). Supporting Interaction in the RoboCare Intelligent Assistive Environment. *Proceedings of AAAI Spring Symposium on Interaction Challenges for Intelligent Assistants*, Stanford, USA: 18-25
Cesta, A. and F. Pecora (2005). The ROBOCARE Project: Intelligent Systems for Elder Care. *Proceedings of the AAAI Fall Symposium on "Caring Machines: AI in Elder Care"*, Washington DC, USA: 1-4
Chartrand, T. L. and J. A. Bargh (1999). The chameleon effect: the perception-behavior link and social interaction. *Journal of Personality and Social Psycholology* 76(6): 893-910.
Chau, P. Y. K. (2001). Influence of Computer Attitude and Self-efficacy on IT Usage Behavior. *Journal of End User Computing* 13(1): 26-33.
Chesney, T. (2006). An Acceptance Model for Useful and Fun Information Systems. *Human Technology* 2(2): 225-235.
Cheverst, K., H. E. Byun, D. Fitton, C. Sas, C. Kray and N. Villar (2005). Exploring issues of user model transparency and proactive behaviour in an office environment control system. *User Modeling and User-Adapted Interaction* 15(3): 235-273.
Cho, K., W. Jung, S. Moon, S. Park and C. Ko (2004). Life guidelines of the Sasang constitutional medicine in the management of'Mibyou'in Korea. *Geriatrics and Gerontology International* 4(s1): 216-219.
Chou, C. P. and P. M. Bentler (1990). Model modification in covariance structure modeling: A comparison among likelihood ratio, Lagrange Multiplier, and Wald tests. *Multivariate Behavioral Research* 25(1): 115-136.
Chow, G. C. (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica* 28(3): 591-605.
Clarkson, E. and R. C. Arkin (2007). Applying Heuristic Evaluation to Human-Robot Interaction Systems. *Proceedings FLAIRS Conference*, Key West, USA: 44-49
Cody-Allen, E. and R. Kishore (2006). An extension of the UTAUT model with e-quality, trust, and satisfaction constructs. *Proceedings of the 2006 ACM SIGMIS CPR conference on computer personnel research*, Claremont, California, USA: 82-89
Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement* 20(1): 37-46.
Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science* 1(3): 98-101.
Coker, D. A. and J. Burgoon (1987). The Nature of Conversational Involvement and Nonverbal Encoding Patterns. *Human Communication Research* 13(4): 463-494.
Compeau, D. and C. Higgins (1995). Computer Self Efficacy: Development of a Measure and Initial Test. *Management Information Systems Quarterly* 19(1): 189-211
Conati, C., S. Marsella and A. Paiva (2005). Affective interactions: the computer in the affective loop. *Proceedings of the 10th international conference on Intelligent user interfaces*, San Diego, California, USA: 1-8
Consolvo, S., P. Roessler, B. E. Shelton, A. LaMarca and B. Schilit (2004). Technology for Care Networks of Elders. *Pervasive computing* 3(2): 22-29

Cowell, A. J. and K. M. Stanney (2005). Manipulation of non-verbal interaction style and demographic embodiment to increase anthropomorphic computer character credibility. *International Journal of Human-Computer Studies* 62: 281-306.
Cramér, H. (1999). *Mathematical Methods of Statistics*. Princeton, New Jersey, USA, Princeton University Press.
Dahlbäck, N., A. Jönsson and L. Ahrenberg (1993). Wizard of Oz studies: why and how, ACM Press New York, NY, USA.
Dario, P. (1999). MOVAID: a personal robot in everyday life of disabled and elderly people. *Technology and Disability* 10(2): 77-93.
Dautenhahn, K. (2004). Robots we like to live with?! - a developmental perspective on a personalized, life-long robot companion. *Invited paper, Proceedings IEEE Ro-man*, Kurashiki, Japan: 17-22
Dautenhahn, K. and C. L. Nehaniv (2002). An agent-based perspective on imitation. *Imitation in Animals and Artifacts*. K. Dautenhahn and C. L. Nehaniv. Cambridge, Massachusetts, USA MIT Press**:** 1-40.
Davis, F. D. (1986). *Technology Acceptance Model for Empirically Testing New End-user Information Systems Theory and Results*. Unpublished Doctoral Dissertation, MIT.
Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* September(13): 319-340
Davis, F. D. (1993). User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *International Journal of Man-Machine Studies* 38(3): 475-487
Davis, F. D., R. P. Bagozzi and P. R. Warshaw (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management Science* 35(8): 982-1003.
Davis, F. D., R. P. Bagozzi and P. R. Warshaw (1992). Extrinsic and intrinsic motivation to use computers in the workplace. *Journal of Applied Social Psychology* 22: 1111-1132.
De Ruyter, B., P. Saini, P. Markopoulos and A. J. N. van Breemen (2005). Assessing the Effects of Building Social Intelligence in a Robotic Interface for the Home. *Special Issue of IwC: social impact of emerging technologies* 17(5): 522-541.
Decker, M. (2008). Caregiving robots and ethical reflection: the perspective of interdisciplinary technology assessment. *AI & Society* 22(3): 315-330.
Decoster, J. and H. M. Claypool (2004). A Meta-Analysis of Priming Effects on Impression Formation Supporting a General Model of Informational Biases. *Personality and social psychology review* 8(1): 2-27
Dillon, A. (2001). User acceptance of information technology. *Encyclopedia of Human Factors and Ergonomics*. W. Karwowski. London, Taylor and Francis.
Ding, L., W. F. Velicer and L. L. Harlow (1995). Effects of estimation methods, number of indicators per factor, and improper solutions on structural equation modeling fit indices. *Structural Equation Modeling* 2(2): 119-143.
DiSalvo, C., F. Gemperle, J. Forlizzi, E. Montgomery, W. Yonkers and J. Divine (2003). The Hug: An Exploration of Robotic Form for Intimate Communication. *Proceedings RO-MAN*, Milbrae, USA: 403-408
DiSalvo, C. F., F. Gemperle, J. Forlizzi and S. Kiesler (2002). All robots are not created equal: the design and perception of humanoid robot heads. *Proceedings of the conference on Designing interactive systems: processes, practices, methods, and techniques*, New York, USA: 321-326
Ebersole, P., P. A. Hess and A. S. Luggen (2003). *Toward healthy aging: Human needs and nursing response*. Philadelphia, Pennsylvania, WB Saunders.
Elliott, C., J. Rickel and J. Lester (1997). Integrating affective computing into animated tutoring agents. *Proceedings of the IJCAI-97 Workshop on Animated Interface Agents*, Nagoya, Japan: 113-121
Evans, I. M. (1992). Peer Interactions and Social Acceptance of Elementary-Age Children with Severe Disabilities in an Inclusive School. *Journal of the Association for Persons with Severe Handicaps (JASH)* 17(4): 205-212.

Feil-Seifer, D. and M. J. Mataric (2005). Defining Socially Assistive Robotics. *9th International Conference on Rehabilitation Robotics (ICORR) 2005*: 465-468

Feil-Seifer, D., K. Skinner and M. J. Mataric (2007). Benchmarks for evaluating socially assistive robotics. *Interaction Studies* 8(3): 423-439.

Fink, J. and A. Kobsa (2002). User modeling for personalized city tours. *Artificial intelligence review* 18(1): 33-74.

Fishbein, M. (1980). A theory of reasoned action: some applications and implications. *Proceedings Nebraska symposium on motivation*, Lincoln, Nebraska: 65-116

Fishbein, M. and I. Ajzen (1975). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Reading, Massachusetts, Addison-Wesley Pub. Co.

Fong, T., I. Nourbakhsh and K. Dautenhahn (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems* 42: 143-166.

Fong, T., I. Nourbakhsh, C. Kunz, L. Fluckiger and J. Schreiner (2005). The Peer-to-Peer Human-Robot Interaction Project. *AIAA Space*, Long Beach, USA: 1-11

Forlizzi, J. (2005). Robotic products to assist the aging population. *Interactions, Volume 12 Issue 2* march: 16-18.

Forlizzi, J. (2007). How robotic products become social products: an ethnographic study of cleaning in the home. *ACM SIGCHI/SIGART Human-Robot Interaction*: 129-136.

Forlizzi, J., C. DiSalvo and F. Gemperle (2004). Assistive Robotics and an Ecology of Elders Living Independently in Their Homes. *Journal of HCI Special Issue on Human-Robot Interaction* 19(1/2): 25 - 59.

Foster, M. E., E. G. Bard, M. Guhe, R. L. Hill, J. Oberlander and A. Knoll (2008). The roles of haptic-ostensive referring expressions in cooperative, task-based human-robot dialogue. *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, Amsterdam, The Netherlands: 295-302

Frambach, R. T. and N. Schillewaert (2002). Organizational innovation adoption: a multi-level framework of determinants and opportunities for future research. *Journal of Business Research* 55(2): 163-176.

Francony, J. M., E. Kuijpers and Y. Polity (1992). Towards a methodology for wizard of oz experiments. *Third Conference on Applied Natural Language Processing*, Trento, Italy.

Friedman, B., P. Kahn and J. Hagman (2003). Hardware companions?: what online AIBO discussion forums reveal about the human-robotic relationship. *Proceedings of the SIGCHI conference on Human factors in computing systems*, Fort Lauderdale, USA.

Fujita, M. (2001). AIBO: Toward the Era of Digital Creatures. *The International Journal of Robotics Research* 20(10): 273-280.

Fujita, M. (2004). On Activating Human Communications With Pet-Type Robot AIBO. *Proceedings of the IEEE* 92(11): 1804-1813.

Gallivan, M. J. (2001). Organizational adoption and assimilation of complex technological innovations: development and application of a new framework. *ACM Sigmis Database* 32(3): 51-85.

Games, P. A. and J. F. Howell (1976). Pairwise multiple comparison procedures with unequal n's and/or variances: a Monte Carlo study. *Journal of Educational and Behavioral Statistics* 1(2): 113-125

Gefen, D., D. Straub and M. Boudreau (2000). Structural equation modeling and regression: Guidelines for research practice. *Structural Equation Modeling* 4(7): 1-78.

Giles, H. and P. F. Powesland (1975). *Speech style and social evaluation*. New York, USA, Academic Press.

Gillies, M. and D. Ballin (2004). Integrating Autonomous Behavior and User Control for Believable Agents. *Third International Joint Conference on Autonomous Agents and Multiagent Systems*, New York, USA: 336-343

Giuliani, M. V., M. Scopelliti and F. Fornara (2005). Elderly people at home: technological help in everyday activities. *Proceedings Ro-man 2005*, New York, USA: 365-370

Gockley, R., J. Forlizzi and R. Simmons (2007). Natural person-following behavior for social robots. *Proceedings of the ACM/IEEE international conference on Human-robot interaction*, Arlington, Virginia, USA: 17-24

Gockley, R., R. Simmons and J. Forlizzi (2006). Modeling Affect in Socially Interactive Robots. *Proceedings Ro-man*, Hertforshire, Hatfield, UK: 558-563

Gomi, T. and A. Griffith (1998). Developing intelligent wheelchairs for the handicapped. *Assistive technology and AI* 1458: 150-178.

Goodhue, D. L. (1995). Understanding User Evaluations of Information Systems. *Mamagement Science* 41: 1827-1844.

Goodhue, D. L. and R. L. Thompson (1995). Task-technology fit and individual performance. *MIS Quarterly* 19(2): 213-236.

Goodrich, M. A. and A. C. Schultz (2007). Human–Robot Interaction: A Survey. *Human–Computer Interaction* 1(3): 203-275.

Gopal, A., S. M. Miranda, B. P. Robichaux and R. P. Bostrom (1997). Leveraging Diversity with Information Technology: Gender, Attitude, and Intervening Influences in The Use of Group Support Systems. *Small Group Research* 28(1): 29-71.

Graf, B. (2001). Reactive navigation of an intelligent robotic walking aid. *Proceedings Ro-man*, Bordeaux-Paris, France: 353-358

Graf, B., M. Hans and R. D. Schraft (2004). Care-O-bot II—Development of a Next Generation Robotic Home Assistant. *Autonomous Robots* 16(2): 193-205.

Grandison, T. and M. Sloman (2000). A survey of trust in internet applications. *IEEE Communications Surveys and Tutorials* 3(4): 2-16.

Green, P. and L. Wei-Haas (1985). The Rapid Development of User Interfaces: Experience with the Wizard of Oz Method. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, Santa Monica, California, USA: 470-474

Greene, K., J. L. Hale and D. L. Rubin (1997). A Test of the Theory of Reasoned Action in the Context of Condom Use and AIDS. *Communication reports* 10: 21-34.

Gresham, F. M. and S. N. Elliot (1990). *Social skills rating system*, Circle Pines: American Guidance Service.

Gresham, F. M., D. L. MacMillan and K. M. Bocian (1998). Agreement between school study team decisions and authoritative definitions in classification of students at-risk for mild disabilities. *School Psychology Quarterly* 13: 181-191.

Guizzo, E. and H. Goldstein (2005). The rise of the body bots [robotic exoskeletons]. *IEEE Spectrum* 42(10): 50-56.

Gumussoy, C. A., F. Calisir and A. Bayram (2007). Understanding the behavioral intention to use ERP systems: An extended technology acceptance model. *Proceedings of the 2007 IEEE IEEM*, Singapore: 2024-2028

Hair, J. F., R. E. Anderson, R. L. Tatham and W. C. Black (1998). *Multivariate Data Analysis with Readings*. Upper Saddle River, New Jersey, Printice Hall International.

Harter, S. and R. Pike (1984). The pictorial scale of perceived competence and social acceptance for young children. *Child Development* 55(6): 1969-1982.

Haslett, B. (1990). Social class, social status and communicative behavior. *Handbook of language and social psychology*: 329–344.

Hayduk, L. A. (1987). *Structural Equation Modeling with LISREL: Essentials and Advances*. Baltimore, Maryland, USA, Johns Hopkins University Press.

Heerink, M., B. J. A. Kröse, B. J. Wielinga and V. Evers (2006a). The Influence of a Robot's Social Abilities on Acceptance by Elderly Users. *Proceedings RO-MAN*, Hertfordshire, UK: 430-439

Heerink, M., B. Kröse, V. Evers and B. Wielinga (2006b). Studying the acceptance of a robotic agent by elderly users. *International Journal of Assistive Robotics and Mechatronics* 7(3): 33-43.

Heerink, M., B. Kröse, B. Wielinga and V. Evers (2006c). Human-Robot User Studies in Eldercare: Lessons Learned. *Smart Homes And Beyond: ICOST2006: 4th International Conference on Smart Homes and Health Telematics*, Belfast, UK: 31-38

Heerink, M., B. J. A. Kröse, B. J. Wielinga and V. Evers (2008a). Enjoyment, Intention to Use and Actual Use of a Conversational Robot by Elderly People. *Proceedings of the third ACM/IEEE International Conference on Human-Robot Interaction*, Amsterdam, The Netherlands.

Heerink, M., B. J. A. Kröse, B. J. Wielinga and V. Evers (2008b). The Influence of Social Presence on Acceptance of a Companion Robot by Older People. *Journal of Physical Agents – Special Issue on Human interaction with domestic robots* 2(2): 33-40.

Heerink, M., B. J. A. Kröse, B. J. Wielinga and V. Evers (2008c). The Influence Of Perceived Adaptiveness Of A Social Agent On Acceptance By Elderly Users. *ISG 2008 - The 6th International Conference of the Internat ional Society for Gerontechnology*, Pisa.

Heerink, M., B. J. A. Kröse, B. J. Wielinga and V. Evers (2008d). Measuring Perceived Adaptiveness in a robotic eldercare companion. *HRI 2008 - Workshop Robotic Helpers*, Amsterdam, The Netherlands.

Heerink, M., B. J. A. Kröse, B. J. Wielinga and V. Evers (2008e). The Influence of Social Presence on Enjoyment and Intention to Use of a Robot and Screen Agent by Elderly Users. *Proceedings Ro-man*, München, Germany: 695-700

Heerink, M., B. Kröse, B. Wielinga and V. Evers (2009a). Measuring the influence of social abilities on acceptance of an interface robot and a screen agent by elderly users. *HCI 2009*, Cambridge, UK: 430-439

Heerink, M., B. Kröse, B. Wielinga and V. Evers (2009c). Influence of Social Presence on Acceptance of an Assistive Social Robot and Screen Agent by Elderly Users. *Advanced Robotics* 23(14): 1909-1923.

Heerink, M., B. J. A. Kröse, V. Evers and B. J. Wielinga (2010a). Relating conversational expressiveness to social presence and acceptance of an assistive social robot *Virtual Reality*, 14(1): 77-84

Heerink, M., B. J. A. Kröse, B. J. Wielinga and V. Evers (2010b). Measuring acceptance of assistive social agent technology by older adults: the Almere model. *International Journal of Social Robotics* 2(3): 1-15

Hennington, A. H. and B. D. Janz (2007). Information systems and healthcare XVI: physician adoption of electronic medical records: applying the UTAUT model in a healthcare context. *Communications of the Association for Information Systems* 19(1/5): 60-80.

Heylen, D., A. Nijholt and D. Reidsma (2006). Determining what people feel and think when interacting with humans and machines: notes on corpus collection and annotation. *Proceedings 1st California Conference on Recent Advances in Engineering Mechanics*, Fullerton, USA: 1-6

Höök, K. (1998). Evaluating the utility and usability of an adaptive hypermedia system. *Knowledge-Based Systems* 10(5): 311-319.

Höök, K., P. Persson and M. Sjölinder (2000). Evaluating users' experience of a character-enhanced information space. *AI Communications* 13(3): 195-212.

Hurst, A., J. Zimmerman, C. Atkeson and J. Forlizzi (2006). The Sense Lounger: Establishing a Ubicomp Beachhead in Elders' Homes. *Proceedings of CHI '06*, Quebec, Canada: 1467-1470

Jacobsson, M., J. Bodin and L. E. Holmquist (2008). The see-Puck: a platform for exploring human-robot relationships. *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, Florence, Italy: 141-144

Jameson, A. (2003). Adaptive interfaces and agents. *The human-computer interaction handbook: fundamentals, evolving technologies, and emerging applications*. J. A. Jacko and A. Sears. Mahwah, Lawrence Erlbaum Assoc Inc**:** 305-330.

Jorge, J. A. (2001). Adaptive tools for the elderly: new devices to cope with age-induced cognitive disabilities. *Proceedings of the 2001 EC/NSF workshop on Universal accessibility of ubiquitous computing*, Alcácer do Sal, Portugal: 66-70

Jung, J. W., J. H. Do, Y. M. Kim, K. S. Suh, D. J. Kim and Z. Z. Bien (2005). Advanced Robotic Residence for the Elderly/the Handicapped: Realization and User Evaluation.

Proceedings of the 9th International Conference on Rehabilitation Robotics, Chicago, USA: 492-495

Kahn, P. H., N. G. Freier, B. Friedman, R. L. Severson and E. Feldman (2004). Social and moral relationships with robotic others? in *Proceedings of the 13th International Workshop on Robot and Human Interactive Communication*: 545-550

Kahn, P. H., H. Ishiguro, B. Friedman and T. Kanda (2006). What is a Human? – Toward Psychological Benchmarks in the Field of Human-Robot Interaction. *Proceedings RO-MAN*, Hatfield, UK: 364-371

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and psychological measurement* 20(1): 141-151.

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika* 35(4): 401-415.

Kanamori, M., M. Suzuki, H. Oshiro, M. Tanaka, T. Inoguchi, H. Takasugi, Y. Saito and T. Yokoyama (2003). Pilot study on improvement of quality of life among elderly using a pettype robot. *Proceedings IEEE International Symposium on Computational Intelligence in Robotics and Automation*, Kobe, Japan: 107-112

Kaplan, F. (2004). Who is afraid of the humanoid? Investigating cultural differences in the acceptance of robots. *International Journal of Humanoid Robotics* 1(3): 1-16

Kawamura, K., P. Nilas, K. Muguruma, J. A. Adams and C. Zhou (2003). An Agent-Based Architecture for an Adaptive Human-Robot Interface. *Proceedings 36th Annual International Conference on System Sciences*, Hawaii: 126-133

Kazerooni, H. (2005). Exoskeletons for human power augmentation. *Proceedings Intelligent Robots and Systems (IROS) Conference*, Edmonton, Alberta, Canada: 3459 - 3464

Kidd, C. and C. Breazeal (2005). Sociable Robot Systems for Real-World Problems. *14th IEEE International Workshop on Robot and Human Interactive Communication*, Nashville, USA: 353-358

Kidd, C. D., W. Taggart and S. Turkle (2006). A Sociable Robot to Encourage Social Interaction among the Elderly. *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*: 3972-3976.

Kihlstrom, J. F. and N. Cantor (2000). Social intelligence. *Handbook of intelligence, 2nd ed.* J. Sternberg. Cambridge, U.K, Cambridge University Press**:** 359-379.

Knutsen, L. A. (2005). M-service expectancies and attitudes: Linkages and effects of first impressions. *Proceedings of the 38th International Conference on System Sciences*, Hawaii: 84-93

Konstantopoulos, S., V. Karkaletsis and C. Matheson (2008). Robot personality: Representation and externalization. *Proceedings Computational Aspects of Affective and Emotional Interaction, Patras, Greece*: 5-13

Kriglstein, S. and G. Wallner (2005). HOMIE: an artificial companion for elderly people. *Conference on Human Factors in Computing Systems*: 2094-2098.

Kruijff, G. J., P. Lison, T. Benjamin, H. Jacobsson, H. Zender, I. Kruijff-Korbayová and N. Hawes (2009). Situated dialogue processing for human-robot interaction. *Cognitive Systems* 8: 311-364.

Kulic, D. and E. A. Croft (2007). Affective State Estimation for Human–Robot Interaction. *IEEE Transactions on Robotics* 23(5): 991-1000.

Landis, J. R. and G. G. Koch (1977). The measurement of observer agreement for categorical data. *Biometrics* 33(1): 159-174.

Lee, B., G. D. Hope and N. J. Witts (2006). Could next generation androids get emotionally close? 'Relational closeness' from human dyadic interactions. *Proceedings RO-MAN*, Hatfield, UK: 475-479

Lee, K. M. and C. Nass (2003). Designing social presence of social actors in human computer interaction. *Proceedings of the SIGCHI conference on Human factors in computing systems*, Fort Lauderdale, USA: 289-296

Lee, V. and H. Wagner (2002). The Effect of Social Presence on the Facial and Verbal Expression of Emotion and the Interrelationships Among Emotion Components. *Journal of Nonverbal Behavior* 26(1): 3-25.

Lee, Y., K. A. Kozar and K. R. T. Larsen (2003). The technology acceptance model: Past, present, and future. *Communications of the Association for Information Systems* 12(50): 752-780

Leifer, L., G. Toye and M. Van Der Loos (1996). Tele-service-robot: Integrating the socio-technical framework of human service through the InterNet-world-wide-web. *Robotics and autonomous systems* 18(1-2): 117-126.

Leite, I. (2007). *iCat, the chess tutor: an affective game buddy based on anticipatory mechanisms*, MSc Thesis, Instituto Superior Técnico.

Leonard-Barton, D. and I. Deschamps (1988). Managerial influence in the implementation of new technology. *Management Science* 34(10): 1252-1265.

Li, J. P. and R. Kishore (2006). How robust is the UTAUT instrument?: a multigroup invariance analysis in the context of acceptance and use of online community weblog systems. *Proceedings of the 2006 ACM SIGMIS CPR conference on computer personnel research*, Claremont, USA: 183-189

Libin, A. and J. Cohen-Mansfield (2004). Therapeutic robocat for nursing home residents with dementia: preliminary inquiry. *American Journal of Alzheimer's Disease and Other Dementias* 19(2): 111-116

Lin, L. I. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45(1): 255-268

Lin, L. I. K. (2000). A note on the concordance correlation coefficient. *Biometrics* 56(1): 324-325.

Lippert, S. K. and M. Davis (2006). A conceptual model integrating trust into planned change activities to enhance technology adoption behavior. *Journal of Information Science* 32(5): 434-448.

Lombard, M. and T. B. Ditton (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated-Communication* 3(2): Available online: http://www.ascusc.org/jcmc/vol3/issue2/lombard.html.

Looije, R., F. Cnossen and M. A. Neerincx (2006). Incorporating guidelines for health assistance into a socially intelligent robot. *Proceedings Ro-man*, Hatfield, UK: 515-520

Louho, R., M. Kallioja and P. Oittinen (2006). Factors Affecting the Use of Hybrid Media Applications. *Graphic Arts in Finland* 35(3): 11-21.

Macer, D. and M. A. C. Ng (2000). Changing attitudes to biotechnology in Japan. *Nature Biotechnology* 18(9): 945-947.

Maciuszek, D. and N. Shahmehri (2003). A framework for the specification of multifunctional, adaptive, and realistic virtual companions for later life. *Proceedings of the International Conference on Aging, Disability and Independence (ICADI)*, Washington DC, USA: 1-15

Maes, P. (1994). Agents that reduce work and information overload. *Communications of the ACM* 37(7): 30-40.

Malhotra, Y. and D. Galletta (1999). Extending the Technology Acceptance Model to Account for Social Influence: Theoretical Bases and Empirical Validation. *Thirty-Second Annual International Conference on System Sciences*, Hawaii: 1006-1019

Marble, J., L. David, J. Bruemmer, D. A. Few and D. D. Dudenhoeffer (2004). Evaluation of Supervisory vs. Peer-Peer Interaction with Human-Robot Teams. *Proceedings of the 37th Annual Hawaii International Conference on System Sciences*

Markopoulos, P., B. de Ruyter, P. Saini and A. van Breemen (2005). Bringing Social Intelligence Into Home Dialogue Systems. *Interactions* july/august: 37-44.

Marsh, S., P. Briggs and W. Wagealla (2004). Considering trust in ambient societies. *Extended abstracts of the 2004 conference on Human factors and computing systems*, Vienna, Austria: 1707-1708

Mathieson, K., E. Peacock and W. W. Chin (2001). Extending the Technology Acceptance Model: The Influence of Perceived User Resources. *The DATA BASE for Advances in Information Systems* 32(3): 86 - 112.

McClelland, M. M. and F. J. Morrison (2003). The emergence of learning-related social skills in preschool children. *Early Childhood Research Quarterly* 18(2): 206-224.

McFarland, D. J. and D. Hamilton (2006). Adding contextual specificity to the technology acceptance model. *Computers in Human Behavior* 22(3): 427-447.

Miller, C. A., P. Wu, K. Krichbaum and L. Kiff (2004). Automated Elder Home Care: Long Term Adaptive Aiding and Support We Can Live With. *Proceedings AAAI Spring Symposium*, Stanford, USA

Mitchell, R. J. (1992). Testing evolutionary and ecological hypotheses using path analysis and structural equation modelling. *Functional Ecology* 6: 123-129.

Mitsunaga, N., Z. Miyashita, K. Shinozawa, T. Miyashita, H. Ishiguro and N. Hagita (2008). What makes people accept a robot in a social environment-discussion from six-week study in an office. *IROS*, Nice, France: 3336-3343

Mival, O., S. Cringean and D. Benyon (2004). Personification Technologies: Developing Artificial Companions for Older People. *Proceedings CHI '04*, Vienna, Austria: 1-8

Montazemi, A. R., D. A. Cameron and K. M. Gupta (1996). An Empirical Study of Factors Affecting Software Package Selection. *Journal of Management Information Systems* 13(1): 89-105.

Montemerlo, M., J. Pineau, N. Roy, S. Thrun and V. Verma (2002). Experiences with a mobile robotic guide for the elderly. *Proceedings of the AAAI National Conference on Artificial Intelligence*, Edmonton, Canada: 587-592

Montgomery, D. C., E. A. Peck and G. G. Vining (2001). Introduction to linear regression analysis. *Wiley series in probability and mathematical statistics,* Wiley Inc. Hoboken,USA.

Moore, G. C. and I. Benbasat (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research* 2(3): 192-222.

Morris, M. G. and V. Venkatesh (2000). Age differences in technology adoption decisions: Implications for a changing work force. *Personnel Psychology* 53(2): 375-404.

Moshkina, L. and R. C. Arkin (2005). Human perspective on affective robotic behavior: A longitudinal study, in Proceedings International Conference on Intelligent Robots and Systems, IROS '05: 1444-1451

Mukai, T., M. Onishi, T. Odashima, S. Hirano and Z. Luo (2008). Development of the Tactile Sensor System of a Human-Interactive Robot "RI-MAN". *IEEE Transactions on Robotics* 24(2): 505-512.

Mynatt, E. D., I. Essa and W. Rogers (2000). Increasing the opportunities for aging in place. *Proceedings of the CUU 2000 Conference on Universal Usability*, New York, USA: 65-71

Nakano, Y. I. and T. Nishida (2005). Awareness of Perceived World and Conversational Engagement by Conversational Agents. *AISB 2005 Symposium: Conversational Informatics for Supporting Social Intelligence & Interaction*, Hatfield, UK

Nijholt, A. (2003). Disappearing Computers, Social Actors and Embodied Agents. *2003 International Conference on Cyberworlds*, Singapore: 128-134

Nomura, T., T. Suzuki, T. Kanda and K. Kato (2006). Measurement of Anxiety toward Robots. *Proceedings RO-MAN '06,* Hatfield, UK: 372-377

Nourbakhsh, I. R., J. Bobenage, S. Grange, R. Lutz, R. Meyer and A. Soto (1999). An affective mobile robot educator with a full-time job. *Artificial Intelligence* 114(1-2): 95-124.

Nunnaly, J. C. and I. H. Bernstein (1978). *Psychometric theory*. New York, McGraw-Hill

Orlikowski, W. J. (1992). The duality of technology: Rethinking the concept of technology in organizations. *Organization science* 3(3): 398-427.

Orlikowski, W. J. (2000). *Using technology and constituting structures: A practice lens for studying technology in organizations*. Organization Science (11:4): 404-428.

Pacchierotti, E., H. I. Christensen and P. Jensfelt (2006). Evaluation of passing distance for social robots. *Proceedings Ro-man '06*, Hatfield, UK: 315-320

Park, K. H., H. E. Lee, Y. Kim and Z. Z. Bien (2008). A Steward Robot for Human-Friendly Human-Machine Interaction in a Smart House Environment. *IEEE Transactions on Automation Science and Engineering* 5(1): 21-25.

Pereira, A., C. Martinho, I. Leite and A. Paiva (2008). iCat, the chess player: the influence of embodiment in the enjoyment of a game. *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems*, Estoril, Portugal: 1253-1256

Pew, R. and S. V. Hemel (2004). *Technology for Adaptive Aging*. Washington, D.C., National Academy Press.

Picard, R. W. (1997). *Affective Computing*. Cambridge, Massachusetts, MIT Press.

Picard, R. W. and S. B. Daily (2005). Evaluating affective interactions: Alternatives to asking what users feel. *Presented at the 2005 CHI Workshop 'Evaluating Affective Interfaces'*, Portland, Oregon, USA.

Pineau, J., M. Montemerlo, M. Pollack, N. Roy and S. Thrun (2003). Towards robotic assistants in nursing homes: Challenges and results. *Robotics and Autonomous Systems* 42: 271-281.

Pineau, J., M. Montemerlo, M. Pollack, N. Roy and S. Thrun (2003). Towards robotic assistants in nursing homes: Challenges and results. *Robotics and Autonomous Systems* 42(3-4): 271-281.

Pollack, M. (2005). Intelligent Technology for an Aging Population: The Use of AI to Assist Elders with Cognitive Impairment. *AI Magazine* 26(2): 9-24.

Pollack, M. E., S. Engberg, J. T. Matthews, S. Thrun, L. Brown, D. Colbry, C. Orosz, B. Peintner, S. Ramakrishnan and J. Dunbar-Jacob (2002). Pearl: A mobile robotic assistant for the elderly. *AAAI Workshop on Automation as Eldercare* 2002: 85-92

Price, B. A., K. Adam and B. Nuseibeh (2005). Keeping ubiquitous computing to yourself: A practical model for user control of privacy. *International Journal of Human-Computer Studies* 63(1-2): 228-253.

Pynoo, B., P. Devolder, T. Voet, J. Vercruysse, L. Adang and P. Duyck (2007). Attitude as a Measure for Acceptance: Monitoring IS Implementation in a Hospital Setting. *Proceedings SIGHCI*, Montréal, Canada: 20-24

Reeves, B. and C. Nass (1996). *The Media Equation: How People Treat Computers, Televisions, and New Media as Real People and Places*. New York, Cambridge University Press.

Rogers, E. M. (1995). *Diffusion of Innovations*. New York, Free Press.

Russell, C. L., V. S. Conn and P. Jantarakupt (2006). Older adult medication compliance: integrated review of randomized controlled trials. *American Journal of Health Behavior* 30(6): 636-650.

Saito, T., T. Shibata, K. Wada and K. Tanie (2003). Relationship between interaction with the mental commit robot and change of stress reaction of the elderly.. *Proceedings of IEEE CIRA*, Kobe Japan: 16-20.

Santos, J. R. A. (1999). Cronbach’s alpha: A tool for assessing the reliability of scales. *Journal of Extension* 37(2): 1-5.

Satake, S., T. Kanda, D. F. Glas, M. Imai, H. Ishiguro and N. Hagita (2009). How to approach humans? - strategies for social robots to initiate interaction. *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction* La Jolla, California, USA: 109-116

Scanaill, C. N., S. Carew, P. Barralon, N. Noury, D. Lyons and G. M. Lyons (2006). A review of approaches to mobility telemonitoring of the elderly in their living environment. *Annals of biomedical engineering* 34(4): 547-563.

Schaper, L. K. and G. P. Pervan (2007). ICT and OTs: A model of information and communication technology acceptance and utilisation by occupational therapists. *International journal of medical informatics* 76: 212-221.

Scheiner, S. M., R. J. Mitchell and H. S. Callahan (2000). Using path analysis to measure natural selection. *Journal of Evolutionary Biology* 13(3): 423-433.

Scherer, K. R. (1987). Toward a dynamic theory of emotion: The component process model of affective states. *Geneva Studies in Emotion and Communication* 1(1): 1-98.
Scheutz, M., J. Kramer, C. Middendorff, P. Schermerhorn, M. Heilman, D. Anderson and P. Bui (2005). Toward affective cognitive robots for human-robot interaction, Menlo Park, CA; Cambridge, MA; London; AAAI Press
Schmidt, A. (2005). *Interactive Context-Aware Systems Interacting with Ambient Intelligence*. Amsterdam, IOS Press.
Scholtz, J., B. Antonishek and J. Young (2004). Evaluation of a Human-Robot Interface: Development of a Situational Awareness Methodology. *Proceedings of the 37th International Conference on System Sciences*, Hawaii
Schubert, T., F. Friedmann and H. Regenbrecht (1999). Embodied presence in virtual environments. *Visual representations and interpretations*. R. P. I. Neilson. Berlin, Springer-Verlag**:** 268-278.
Schumacker, R. E. and R. G. Lomax (1996). A beginner's guide to structural equation modeling. *Mahwah, New Jersey*.
Sejwacz, D., I. Ajzen and M. Fishbein (1980). Predicting and understanding weight loss: Intentions, behaviors, and outcomes. *Understanding Attitudes and Predicting Social Behavior*. I. Ajzen and M. Fishbein. Englewood Cliffs, Prentice Hall**:** 101-112.
Shahmehri, N. (2001). Intelligent systems and the elderly–problems and possibilities. *Proceedings Conference on Aging, Care and Welfare of Elderly and how IT can improve Quality of Life*, Stockholm.
Sharma, S., R. M. Durand and O. Gur-Arie (1981). Identification and analysis of moderator variables. *Journal of Marketing Research* 18(August): 291- 300.
Shibata, T., K. Wada and K. Tanie (2003). Statistical Analysis and Comparison of Questionnaire Results of Subjective Evaluations of Seal Robot in Japan and U.K. *Proceedings of the 2003 IEEE International Conference on Robotics & Automation*, Taipei, Taiwan: 3152-3157
Shinozawa, K., F. Naya, K. Kogure and J. Yamato (2004). Effect of robot's tracking users on human decision making. *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sendai, Japan: 1908-1913
Shinozawa, K., F. Naya, J. Yamato and K. Kogure (2005). Differences in Effect of Robot and Screen Agent Recommendations on Human Decision-Making. *IJHCS* 62(2): 267-279.
Shinozawa, K., B. Reeves, K. Wise, S. Lim, H. Maldonado and F. Naya (2003). Robots as New Media: A Cross-Cultural Examination of Social and Cognitive Responses to Robotic and On-Screen Agents. *Proceedings of the 53rd Annual Conference of the International Communication Association, Information Systems Division*, San Diego, California, USA: 998-1002
Shneiderman, B. (1997). Direct Manipulation versus Agents: Paths to Predictable, Controllable, and Comprehensible Interfaces. *Software Agents*. J. M. Bradshaw. Menlo Park California, AAAI Press/MIT Press**:** 97-106.
Sidner, C. L. and C. Lee (2005). Engagement During Dialogues with Robots. *AAAI Spring Symposia*, Stanford, California, USA.
Sixsmith, A. J. (2002). New technology for the care of cognitively impaired older people. *Geriatric Psychiatry*. J. R. M. Copeland, M. T. Abou-Saleh and D. G. Blazer. New York, Wiley**:** 685-687.
Soede, A. J., J. C. Vrooman, P. M. Ferraresi and G. Segre (2004). *Unequal Welfare States. Distributive consequences of population aging in six European countries*. Den Haag, SCP.
Sparks, P., R. Shepherd and L. J. Frewer (1995). Assessing and structuring attitudes towards the use of gene technology in food production: The role of perceived ethical obligation. *Basic and Applied Social Psychology* 16: 267-85.
Stanford, V. (2002). Using pervasive computing to deliver elder care. *IEEE Distributed Systems Online* 3(3): 10-13.

Sternberg, R. J., B. E. Conway, J. L. Ketron and M. Bernstein (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology* 41: 37-55.

Stiehl, W. D., J. Lieberman, C. Breazeal, L. Basel, R. Cooper, H. Knight, L. Lalla, A. Maymin and S. Purchase (2006). The huggable: a therapeutic robotic companion for relational, affective touch. *International Conference on Computer Graphics and Interactive Techniques archive, ACM SIGGRAPH,* Boston, USA

Streefkerk, J. W., M. P. van Esch-Bussemakers and M. A. Neerincx (2006). Designing personal attentive user interfaces in the mobile public safety domain. *Computers in Human Behavior* 22(4): 749-770.

Strunk, B. C., P. B. Ginsburg and J. R. Gabel (2001). *Tracking health care costs*. Health Affairs 20(6) (Suppl Web exclusives): W39-W50,

Suga, K., M. Sato, H. Yonezawa, S. Naga and J. Shimizu (2003). Effects of robot-assisted activity on senior citizens-Indicators of HVA, MHPG, and CS concentrations in saliva. *Journal of Analytical Bio-Science* 26(5): 435-440.

Sun, H. and P. Zhang (2006). Causal Relationships between Perceived Enjoyment and Perceived Ease of Use: An Alternative Approach. *Journal of the Association for Information Systems* 7(9): 618-645.

Sun, H. and P. Zhang (2006). The role of moderating factors in user technology acceptance *International Journal of Human-Computer Studies* 64(2): 53-78.

Tabachnick, B. G. and L. S. Fidell (2001). *Using multivariate statistics*. Needham Heights, Massachusetts, USA, Allyn & Bacon Inc.

Taggart, W., S. Turkle and C. Kidd (2005). An interactive robot in a nursing home: Preliminary remarks. *Towards Social Mechanisms of Android Science*. Stresa, Italy, Cognitive Science Society**:** 56-61.

Tamura, T., S. Yonemitsu, A. Itoh, D. Oikawa, A. Kawakami, Y. Higashi, T. Fujimooto and K. Nakajima (2004). Is an Entertainment Robot Useful in the Care of Elderly People With Severe Dementia? *Journals of Gerontology Series A: Biological and Medical Sciences* 59(1): 83-85.

Tanaka, J. S. (1987). "How Big Is Big Enough?": Sample Size and Goodness of Fit in Structural Equation Models with Latent Variables. *Child Development* 58: 134-146.

Tapus, A., M. J. Mataric and B. Scassellati (2007). The grand challenges in socially assistive robotics. *IEEE Robotics And Automation Magazine Special Issue On Grand Challenges In Robotics* 14(1): 35-42.

Taylor, S. and P. A. Todd (1995). Understanding Information Technology Usage: A Test of Competing Models. *INFORMATION SYSTEMS RESEARCH* 6: 144-176.

Thompson, R., C. Higgins and J. Howell (1991). Personal Computing: Toward a Conceptual Model of Utilization. *Management Information Systems Quarterly* 15(1): 7.

Tscheligi, M. and R. Bernhaupt (2004). HCl overviews: Advanced studies and research in information and communication technologies & society: The ICT&S-Center. *Extended abstracts of the 2004 conference on Human factors and computing systems*, Vienna, Austria.

Turkle, S., W. Taggart, C. D. Kidd and O. Dasté (2006). Relational artifacts with children and elders: the complexities of cybercompanionship. *Connection Science* 18(4): 347-361.

Tzeng, J.-Y. (2004). Toward a more civilized design: studying the effects of computers that apologize. *International journal of human-computer studies* 61(3): 319-345.

United Nations (2007). *World Population Prospects: The 2006 Revision-Comprehensive Tables*. New York, United Nations Publications.

Uzoka, F. M. E. (2008). Organisational influences on e-commerce adoption in a developing country context using UTAUT. *International Journal of Business Information Systems* 3(3): 300-316.

Vallerand, R. J. (1997). Toward a Hierarchical Model of Intrinsic and Extrinsic Motivation. *Advances In Experimental Social Psychology* 29: 271-360.

van Breemen, A., X. Yan and B. Meerbeek (2005). iCat: an animated user-interface robot with personality. *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*: 143-144.
van Breemen, A., X. Yan and B. Meerbeek (2005). iCat: an animated user-interface robot with personality. *Proceedings of the fourth international joint conference on autonomous agents and multiagent systems*, New York, USA: 143-144
Van der Heijden, H. (2004). User acceptance of hedonic information systems. *MIS quarterly* 28(4): 695-704.
Vaughn, S., A. Hogan, K. Kouzekanani and S. Shapiro (1990). Peer acceptance, self-perceptions, and social skills of learning disabled students prior to identification. *Journal of Educational Psychology* 82(1): 101–106.
Venkatesh, V. (2000). Determinants of Perceived Ease of Use Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model. *Information Systems Research* 11(4): 342-365.
Venkatesh, V. and F. D. Davis (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science* 46(2): 186-204.
Venkatesh, V., M. G. Morris, G. B. Davis and F. D. Davis (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly* 27(3): 425-478.
Verzijden, D. and J. Fransen (2004). *Vergrijzing in Nederland - onderzoek uitgevoerd in opdracht van de Rijksvoorlichtingsdienst/Publiek en Communicatie ten behoeve van het ministerie van Volksgezondheid, Welzijn en Sport* Veldkamp.
Wada, K. and T. Shibata (2007). Living With Seal Robots—Its Sociopsychological and Physiological Influences on the Elderly at a Care House. *Robotics, IEEE Transactions on [see also Robotics and Automation, IEEE Transactions on]* 23(5): 972-980.
Wada, K., T. Shibata, T. Saito and K. Tanie (2003a). Effects of robot assisted activity to elderly people who stay at a health service facility for the aged. *Proceedings. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003),* Las Vegas, USA 2847–2852
Wada, K., T. Shibata, T. Saito and K. Tanie (2003b). Psychological and Social Effects of Robot Assisted Activity to Elderly People who Stay at a Health Service Facility for the Aged. *Proceedings International Conference on Robotics and Automation*, Las Vegas, Nevada, USA: 3996–4001
Wada, K., T. Shibata, T. Saito and K. Tanie (2003c). Psychological, physiological and social effects to elderly people by robot assisted activity at a health service facility for the aged. *Proceedings IEEE/ASME International Conference on Advanced Intelligent Mechatronics, AIM 2003,* Kobe, Japan
Wada, K., T. Shibata, T. Saito and K. Tanie (2003d). Relationship between familiarity with mental commit robot and psychological effects to elderly people by robot assisted activity. Proceedings *IEEE International Symposium on Computational Intelligence in Robotics and Automation,* Kobe, Japan: 113-118
Wagner, A. R. (2009). *The role of trust and relationships in human-robot social interaction*, Georgia Institute of Technology.
Wagner, H. L. and J. Smith (1991). Facial expression in the presence of friends and strangers. *Journal of Nonverbal Behavior* 15(4): 201-214.
Walters, M. L., D. S. Syrdal, K. Dautenhahn, R. Te Boekhorst and K. L. Koay (2008). Avoiding the uncanny valley: robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion. *Autonomous Robots* 24(2): 159-178.
Wilkes, D. M., A. Alford, R. T. Pack, T. Rogers, I. I. R. A. Peters and K. Kawamura (1997). Toward socially intelligent service robots. *Applied Artificial Intelligence* 12(7-8): 729-766.
Wilson, E. V. and N. K. Lankton (2004). Modeling Patients' Acceptance of Provider Delivered E-Health. *Journal of the American Medical Informatics Association* 11(4).
Witmer, B. G. and M. J. Singer (1998). Measuring Presence in Virtual environments: A Presence Questionnaire. *Presence* 7(3): 225-240.

Woods, S. N., M. L. Walters, K. L. Koay and K. Dautenhahn (2006a). Comparing Human Robot Interaction Scenarios Using Live and Video Based Methods: Towards a Novel Methodological Approach. *9th IEEE International Workshop on Advanced Motion Control*, Hatfield, UK: 27-29

Woods, S. N., M. L. Walters, K. L. Koay and K. Dautenhahn (2006b). Methodical Issues in HRI: a Comparison of Live and Videobased Methods in Robot to Human Approach Direction Trials. *Proceedings Ro-man*, Hatfield, UK: 6-8

Wu, I.-L. and J.-L. Chen (2005). An extension of Trust and TAM model with TPB in the initial adoption of on-line tax: An empirical study *International Journal of Human-Computer Studies* 62(6): 784-808.

Wu, J.-H., S.-C. Wang and L.-M. Lin (2005). What Drives Mobile Health Care? An Empirical Evaluation of Technology Acceptance. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*: 719-729

Wu, P. and C. Miller (2005). Results from a field study: The need for an emotional relationship between the elderly and their assistive technologies. *Foundations ofAugmented Cogniton* 11: 889-898.

Wu, Y. L., Y. H. Tao and P. C. Yang (2007). Using UTAUT to explore the behavior of 3G mobile communication users, *Proceedings IEEE International Conference on Industrial Engineering and Engineering Management*. Singapore: 199-203

Wubs, H. and F. Huysmans (2006). *Snuffelen en Graven*. S. e. C. Planbureau. The Hague, The Netherlands.

Xaverius, P. K. and R. M. Mathews (2004). Evaluating the Impact of Intergenerational Activities on Elders' Engagement and Expressiveness Levels in Two Settings. *Journal of intergenerational relationships* 1(4): 53-69.

Yanagi, H. and S. Tomura (2002). A Pilot Study for Animal-Assisted Therapy using Companion Animal Type Robot (AIBO) in Primary Care Setting. *Japanese Journal of Primary Care* 25(2): 108-114.

Yanco, H. A. (2001). Development and testing of a robotic wheelchair system for outdoor navigation. *Proceedings of the 2001 Conference of the Rehabilitation Engineering and Assistive Technology Society of North America*, Reno, Nevada, USA.

Yanco, H. A. and J. Drury (2004). Classifying human-robot interaction: an updated taxonomy, Proceedings IEEE International Conference on Systems, Man and Cybernetics, The Hague, The Netherlands: 2841 - 2846

Yang, H. and Y. Yoo (2004). It's all about attitude: revisiting the technology acceptance model. *Decision Support Systems* 38(1): 19-31.

Yi, M. Y. and Y. Hwang (2003). Predicting the use of web-based information systems: self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. *International Journal of Human-Computer Studies* 59(4): 431-449.

Yilmazyildiz, S., W. Mattheyses, Y. Patsis and W. Verhelst (2006). Expressive Speech Recognition and Synthesis as Enabling Technologies for Affective Robot-Child Communication. *Lecture Notes in Computer Science* 4261: 1-8.

Yoon, Y., T. Guimaraes and Q. O'Neal (1995). Exploring the factors associated with expert systems success. *MIS Quarterly* 19(1): 83-106.

Yoshikawa, Y., K. Shinozawa, H. Ishiguro, N. Hagita and T. Miyamoto (2006). Responsive robot gaze to interaction partner. *Proceedings of Robotics: Science and systems* Philadelphia, Pennsylvania, USA.

Yu, H., M. Spenko and S. Dubowsky (2003). An Adaptive Shared Control System for an Intelligent Mobility Aid for the Elderly. *Autonomous Robots, Volume 15 Issue 1* July.

Yuanquan, L. I., Q. I. Jiayin and S. H. U. Huaying (2008). Review of Relationships Among Variables in TAM. *Tsinghua Science and Technology* 13(3): 273-278.

Zhu, Y. Y. and Z. M. Mao (2008). Application study of UNIANOVA and linear in SPSS. *Jisuanji Gongcheng yu Sheji(Computer Engineering and Design)* 29(8): 2108-2110.

Zwick, R. (1986). Testing pairwise contrasts in one-way analysis of variance designs. *Psychoneuroendocrinology* 11(3): 253-276.

Appendix A – Overview of assistive social robots for older adults

To learn more about the individual agents used for experiments with elderly users, we will take a closer look at each of them. To describe each robot, robot type (which will be companion robot, service robot or a mixed type), embodiment, tasks and role, developmental background and results of user studies (if available) will be discussed.

The robots described are selected for our review either because they have been specifically developed for older adults or because they have been subject to user studies in which they were used as assistive social robots for older adults.

Paro

Type
Paro is a typical example of a companion type robot.

Embodiment
Its appearance is that of a soft seal robot. Paro has programmable behavior as well as a set of sensors. Sensors include a touch sensor over the complete body, an infra red sensor, stereoscopic vision and hearing. Actuators include eyelids, upper body motors, front paw and hind limb motors. Paro is not mobile. It has been subject to many user studies, both short and long term, in different countries.

Fig. A-1. Paro

Task/role
It is developed to study the effects of 'Animal Assistive Therapy' (Beck and Katcher 1996) for elderly with companion robots. This means it has pet-like features, also in its way of communication: it can make a purring sound and close its eyes to express satisfaction, and soft cries to respond to vocal attention and produce smooth body movements when being patted.

Developmental background
Paro has been developed by the Intelligent Systems Research Institute (ISRI) of the National Institute of Advanced Industrial Science and Technology (AIST) in Japan, and is produced by Intelligent System Co., Ltd..

User studies
Results show that Paro has an effect similar to that of pets. It decreases feelings of loneliness and has a positive influence on both mental and physical health (Wada et al. 2003a; Wada and Shibata 2007).

Huggable

Type
The Huggable (Stiehl et al. 2006) is a companion type robot with a few service type functionalities.

Embodiment
The Huggable has the shape of a Teddy bear, a choice that was deliberate as the Teddy Bear is a symbol of warmth and comfort familiar to many different age groups and cultures. It features a full body 'sensitive skin' for relational affective touch, silent, muscle-like, voice coil actuators, camera's, microphones and an embedded PC with data collection and networking capabilities.

Fig A-2. The Huggable

Task/role
Its purpose is to be a pet replacement, much like Paro, to replace animal assisted therapy. It detects and responds to petting, rubbing, tapping, scratching, and other types of interactions that person usually has with an animal. Since it is also capable of sensor processing, data storage, and data, video and sound transmission, it can also be used to gather information on the behavior of its users.

Developmental background
The Huggable has been developed at MIT Media Lab.

User studies
At this moment there are no published user studies on the Huggable

Aibo

Type
Aibo has been used as a companion robot.

Embodiment
Aibo has the embodiment of a small dog, although, without a furry skin and eyes, it is still clearly a robot,. has programmable behavior, a hard plastic exterior and has a wide set of sensors and actuators. Sensors include a camera, touch sensors, infra red and stereo sound. Actuators include 4 legs, a moveable tail, and a moveable head.

Fig. A-3. Aibo

Aibo is mobile and autonomous. It can find its power supply by itself and it is programmed to play and interact with humans.

Task/role
Aibo has been used extensively in studies with elderly in order to try to assess the effectiveness on the quality of life and symptoms of stress. Aibo could be programmed to have assistive functionalities (for example monitoring or giving directions), but in the above mentioned user studies these studies, Aibo was a typical companion type robot: not assistive and causing effects on users that are usual for pets.

Developmental background
Aibo is an entertainment robot developed and produced by Sony (Fujita 2001). It is currently out of production.

User studies
Several studies show that Aibo, used as a pet replacement, had a positive effect on helath and wellbeing (Yanagi and Tomura 2002; Kanamori et al. 2003; Suga et al. 2003; Fujita 2004; Mival et al. 2004; Tamura et al. 2004; Turkle et al. 2006).

Homie

Type
Homie is a companion type robot, though it has some service type capabilities.

Embodiment
Homie is a dog shaped combination of a social companion and communication device (Kriglstein and Wallner 2005). It can communicate several emotions and can be used to reproduce messages that have been sent by text. The entire "Homie" system consists of four parts: the dog shaped puppet itself, a console, which is plugged to the television, a dog bed, and a bracelet/collar with sensors that can be used to retrieve user data.

Fig A-4. Homie

Task/role
Homie can serve as a pet-like companion, but it also can be used to collect user data and as a communication intermediate.

Developmental background
developed at the University of Vienna, Austria

User studies
At this moment there are no published user studies on Homie.

iCat

Type
The iCat is a basically a service type robot, although it has some companion type capabilities.

Embodiment
It is made of hard plastic and has a cat-like appearance, with movable lips, eyes, eyelids and eyebrows to display different facial expressions to simulate emotional behavior. There is a camera installed in the iCat's nose which can be used for different computer vision capabilities, such as recognizing objects and faces. The iCat's base contains two microphones to record the sounds it hears and a loudspeaker is built in for sound and speech output.

Task/role
It's design aim is to be a research platform for human-robot interaction, possibly in an intelligent home environment. Studies typically investigate how users perceive the iCat as interface to new technology.

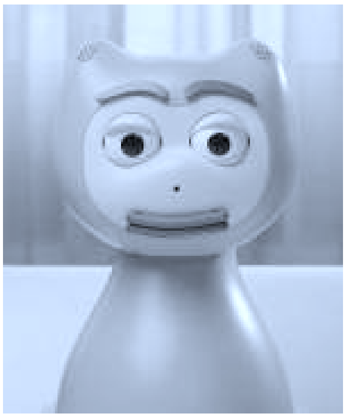


Fig A-5. iCat

Developmental background
The iCat has been developed and is produced by Philips (van Breemen et al. 2005).

User studies
Besides our own studies, iCat has not been used for experiments on elderly users. In an experiment by Looije et al. (2006) it featured as a personal assistant for a small group of people with diabetes. Results show that participants appreciated a more intelligent agent and there would be more chance of them using it than a less social robot.

Nursebot (Flo/Pearl)

Type
The nursebot is a service type robot with some companion type capabilities.

Embodiment
Flo has a head with anthropomorphic features that have been further developed in Pearl. Below the head is a screen that can be used both for (touch screen) input and output of information.

Fig A-6. Flo and Pearl

It is equipped with a differential drive system, two on-board PCs, wireless Ethernet, laser range finders,sonar sensors, microphones for speech recognition, speakers for speech synthesis, touch-sensitive graphical displays, actuated head units, and stereo camera systems. Pearl also features two sturdy handle-bars, a compact design that allows for cargo space, a removable tray, and a sophisticated head unit.

Task/role
It can help elderly to navigate through the nursing facility. It can also remind its users of events and can provide advice and cognitive support.

Developmental background
Flo was the first generation and Pearl is the second generation of nursebots developed by Carnegie Mellon University (Pollack et al. 2002; Pineau et al. 2003).

User studies
Preliminary field tests have been conducted with both older adults and patients with traumatic brain injury, but systematic studies of its effectiveness have not yet been completed.

Care-o-bot

Type
Although Care-o-bot can be used as a companion, it is primarily a service type robot.

Embodiment
Care-o-bot is designed with functional and not anthropomorphic or zoomorphic considerations (see figure 2.9). It has handles that make it possible to be used as a walking aid and versions II and III both have an arm that makes it possible to pick up things.

Task/role
It's task is to provide elderly with physical help. It can serve as a walking aid, a butler and a monitor.

Developmental background
Care-o-bot has been (and is still being) developed by a German consortium headed by Frauenhofer Institute (Graf et al. 2004). At this moment there have been three versions of Care-o-bot.

Fig A-7. Care-o-bot I/II/ III

User studies
Published studies with Care-o-bot are focused on technical development and not on user responses. Remarkable is that experiments wit Car-o-bot were usually not done in a Wizard of Oz setting, meaning that functionalities were not faked. There are however no published studies on the effect it has on health or wellbeing, nor on the tendency to accept this robot.

The RoboCare robot

Type
The RoboCare robot is a service type robot.

Embodiment
The RoboCare robot is cylinder shaped and mobile (wheels). It connects to a system that features sensors and camera's. It is capable of producing pre-programmed speech. There is a version with a screen on which a female face is displayed to embody the conversation.

Fig A-8. Two versions of RoboCare

Task/role
The robot serves both as an interface to the 'smart home' technology and as an autonomous actor, retrieving information from it's intelligent environment and acting upon this.

Developmental background
The RoboCare project of the university of Rome, Italy is not so much focused on developing a robot (Bahadori et al. 2003) as to an environment, an intelligent home of which a robot is an integrated part.

User studies
Published research related to RoboCare is focused on technical matters or design issues – for example comparing responses to a robot with a screen, a face or just a voice. featuring an intelligent home of which a robot is an integrated part .

Intelligent sweet home robots

Type
A screen agent (called *software type robot* by its developers) and *Joy*, a three dimensional robot (called *steward robot* or *hardware type robot* by its developers). Both can be categorized as service type robots. They are categorized by their developers as steward robots (Mukai et al. 2008; Park et al. 2008).

Embodiment

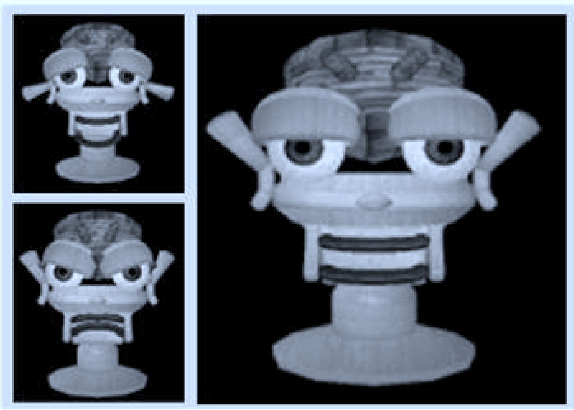


Figure A-9. Screen agent

The software-type robot is represented by a screen agent (a face with humanoid features – see Figure 2.11) and can be accessed everywhere using personal computing devices such as a PDA and a cellular phone when a wired/wireless communication network is available. Facial expression is generated in the emotional interaction module according to the service of the robot and in the context of the interaction.

The hardware-type robot (see Figure 2.12) has a humanoid shape, wit a head that features two eyes.
Both types of robot have an intelligent processing module with learning and emotional interaction capability.

Task/role

Like RoboCare, the Sweet Intelligent Home is a smart environment for elderly. The two robots can function as an interface between this environment and its users (which is mainly the task of the screen agent) and perform services based both on active user input and data form the environment. The hardware-type robot has mainly been developed to provide active services such as delivering a meal or bringing an object with physical interaction by using two robotic arms and a mobile platform.

Developmental background

The steward robots and their smart environment have been developed at the Korea Advanced Institute of Science and Technology (KAIST). The steward robots have been designed (besides on general HRI-pronciples) based on Leifer's design laws for service robots (Leifer et al. 1996). For a human friendly appearance of the hardware agent, the developerd adopted the Sa-sang constitution theory which is the basis of Korean oriental medicine (Cho et al. 2004).

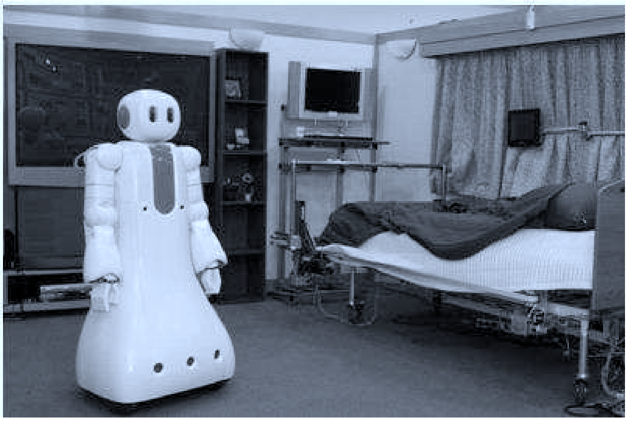


Figure A-10. Joy

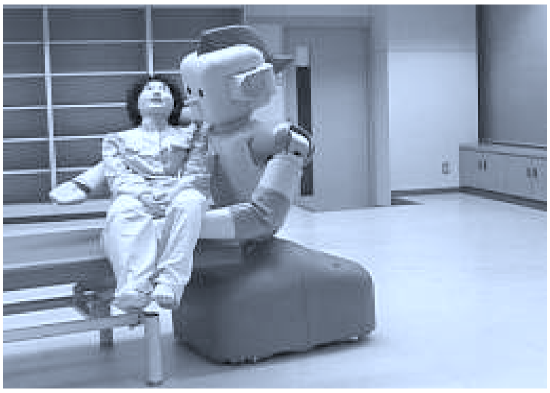
User studies

No user studies have been published yet on these robots or their environments.

RI-MAN

Type
RI-MAN is a service type robot.

Embodiment



RI-MAN is a green 5-foot 2-inch, 220-pound robot with a large wheeled base. The robot has a humanoid appearance, and is equipped with a tactile sensor system that enables it to respond to human touch.

Task/role
It's task is to provide elderly with physical help. It can pick up and deliver objects and bodies (at this point up to 35 kg).

Figure A-11. RI-MAN

Developmental background
RI-MAN the robot, designed at the Bio-mimetic Control Research Center at RIKEN in Nagoya, Japan.

User studies
The robot has demonstrated its ability to carry up and hold a doll safely. It does not yet have the strength to carry real humans.

Laura/FitTrack

Type
Laura is a service type screen agent with companion type features.

Embodiment
Filtrack is a system to stimulate older adults to do physical exercises (Bickmore 2003). Within this system, a female personality called Laura appears as a conversational agent that interacts through messages on the screen. She is capable of many facial expressions, that help her to express care for the user.

Fig A-12. Laura (screenshot)

Task/role
Laura stimulates the user to do exercises and gives instructions on how to carry out these exercises.

Developmental background
The Laura agent is used as a platform to investigate long-term relationships, and the impact of relationship maintenance behaviors on users' reported working alliance with the agent. The appearance and nonverbal behavior of Laura was based on a review of relevant literature and a series of pre-test surveys.

User studies
As we will discuss more detailed in paragraph 2.5, acceptance studies indicate that the agent was accepted by the participants as a conversational partner on health and health behavior and rated high on issues like trust and friendliness. It was also found to be successful as a health advisor (Bickmore 2004; Bickmore et al. 2005).

Annie

Type
Annie is a service type screen agent with companion type features.

Embodiment
Annie is a female humanoid screen character, used in combination with a webcam (attached to the screen), a microphone and two speakers.

Task/role

Fig. A-12. Annie

It's design aim is to be a research platform for human-robot interaction, possibly in an intelligent home environment. Studies typically investigate how users perceive the iCat as interface to new technology.

Developmental background
The screen agent was developed for tests with elderly users in eldercare institutions by students of the Instituut voor Information Engineering in Almere.

User studies
Besides our own studies that will be described in chapter 3, Annie has not been used for experiments on elderly users.

Steffie

Type
Steffie is a service type screen agent with companion type features.

Embodiment

Steffie is designed in Flash. She speaks and uses both facial expressions and gestures. The user communicates with her by clicking buttons that are used for choosing subjects, to let her continue or to let her repeat.

Fig. A-13. Steffie

Task/role

Steffie is developed as a part of a website (www.steffie.nl) where she features as a talking guide for older adults, explaining the internet, e-mail, health insurance, cash dispensers and railway ticket machines.

Developmental background

Steffie has been developed by a consortium of commercial and non-commercial participants, as a part of a project to facilitate the use of the internet by older adults.

User studies

Besides our own studies that will be described in chapter 4, Steffie has not been used for experiments on elderly users.

Appendix B – Overview user studies concerning assistive social robots

Reference	*Design/ task/n*	*Result/ measures*	*Term of use*	*Remarks*
Aibo				
Kanamori et al. 2002	*2,4/1/3*	*+/1,5*	*?*	*All three cases reported to have decreased stress and loneliness, but research method not clear about other influences of visit during the 20 AIBO sessions.*
Kanamori et al. 2003	*2,4/1/5*	*+/1,3,4, 5*	*7 weeks*	*20 sessions, but it is unclear if the positive effects are due to sessions themselves or due to the robot.*
Mival et al. 2004	*2/1/1012*	*+/6*		*A study to find design criteria for companion robots*
Suga et al. 2002	*2,4/1/23*	*+/1*	*2 months*	*Positive immune system response. Unclear how AIBO is used exactly, difficult to attribute causality.*
Sakairi 2004	*2,4/1/8*	*+/3,5*	*30 minutes*	*No statistics reported, difficult to attribute causality.*
Suga et al. 2003	*2,4/1/15*	*+/1*	*?*	*Design is unclear, difficult to attribute causality.*
Tamura et al. 2004	*2,4/1/?*	*±/3*	*5 minutes intervention*	*Although both toy dog (=control) and AIBO increased activity of demented patients, there was less difference (even less patient activity) in the AIBO case then in the toy dog case, probably due to the fact that AIBO was not perceived as a puppy dog.*
Turkle et al. 2006	*2,4/1/2*	*+/5, 6*	*several months*	*Form and behavior of a robot pet might matter for its acceptance. Two cases with positive results in terms of social interaction with the My Real Baby robot*
Yanagi & Tomura 2002	*2/1/46*	*+/5*	*Several hours*	*Study in a waiting room in clinic, the exact result measure is unclear from abstract*
Care-o-bot				
Graf et al. 2004	*2,4/4,6/6*	*+/5*		*Describes results with walking aid robot and grabber. Elderly are able to work with the robot.*
iCat				
Looije et al. 2006	*3,4/3/6*	*±/6,7*	*< 1 hour*	*Study to investigate guidelines for iCat interface design.*
Homie				
Kriglstein and Wallner 2005	*2,4/1,2,5/2*	*+/3, 6*	*?*	*Ideas about design*
Laura/Fit track				
Bickmore & Picard 2005	*3,4/3/8*	*+/6,7*	*?*	
Paro				
Giusti and Marti 2006	*2,4/1/5*	*±/3*	*1 month twice a week*	*Demented elderly started talking a lot about PARO and to PARO, but no clear effect measure. Also difficult to establish causality.*
Kazuyoshi et al. 2003	*2,4/1/12,11*	*±/2*	*3 weeks*	*Same experiment as Saito et al (2003). Different measure. Hints at positive effect, but no statistical analysis. Control (fake PARO) and experimental condition (real PARO) had the same effect.*
Kidd et al. 2006	*1,4/1/23*	*+/3,2*	*4 months*	*More lively communication in PARO-on case compared to PARO-off case. Experimenters took care to not influence sessions. No statistics. Extra result with My Real Baby: is used to calm down residents, but the baby is often too much of a care burden.*
Marti et al. 2006	*2,4/1/1*	*+/3*	*1 time*	*PARO was introduced with therapist. Demented patient accepted PARO and talked about it.*
Saito et al. 2002	*2/1/20*	*+/1*	*6 weeks*	*Urinary tests show a lower stress level*
Saito et al. 2003	*2,4/1/12,11*	*-/5*	*3 weeks, 4 days a week, 1 hour*	*Negative (stress hormone) result for the less active PARO in the less demented group (n=12), but there seems to have been a problem with the urine samples. The more demented group with the active PARO had no results. Again difficult to interpret.*
Taggart et al. 2005	*1,4/1/18*	*+/3*	*20 minutes*	*Form is important for expectations (PARO in bathtub). Acceptation is still an important issue. Less active PARO had fewer reactions of subjects. Result in the opposite direction in study of Saito et al (2003)*

Wada et al. 2002a	*2/1/11*	*±/2*	*3 weeks, 1-3 times per week 20 min*	*Slightly positive results, one item of the mood scale (vigor) was significantly better in Aibo intervention case.*
Wada et al. 2002b	*2/1/11*	*+/2*	*3 weeks, 1-3 times per week 20 min*	*Same as Wada et al (2002)*
Wada et al. 2003a	*2,4/1/4,3,9*	*+/1,2*	*3 weeks, 4 days a week, 1 hour*	*Non-significant increase in immune system function as measured by urinary hormones (n=4). PARO (n=3) and fake PARO group (n=9) both had positive effect on depression*
Wada et al. 2003b	*2,4/1/7,11, 12,9*	*+/2*	*3 weeks, 4 days a week, 1 hour*	*Subjects were happier with the real PARO (n=7) than with fake PARO (n=11) but they kept liking the fake PARO (n=12) better throughout the study compared to the real PARO (n=9).*
Wada et al. 2003c	*2,4/1/4,7,11*	*+/1,2*	*3 weeks*	*Same data and results in Wada (2003b) and (2003c)*
Wada et al. 2003d	*2,4/1/4,7,11*	*+/2*	*3 weeks*	*Correlation between emotion change and familiarity with PARO (n=4). Fake PARO (n=11) has same interest effect as real PARO (n=7), i.e. subjects keep liking both robots.*
Wada et al. 2004a	*2/1/10*	*+/1*	*14 weeks*	*No statistically sound evidence of effect (n=10) of PARO on dementia scale. One case seems promising (woman). Application of PARO to elderly seems different than in other 2003 studies. It is unclear what the amount of involvement of the researcher is.*
Wada et al. 2004b	*2/1/12,11*	*+/2*	*3 weeks, 1-3 times per week 20 min*	*Increase in mood and emotion faces test, but unclear where the effect came from. The intervention with PARO is mediated, and the measurement is before and after intervention.*
Wada et al. 2004c	*2/1/12,11*	*+/2*	*5 weeks, etc.*	*Same as Wada et al (2004c)*
Wada et al. 2005a	*2/1/23*	*+/2*	*1 year*	*Longer term study with few subjects (n=8). Unclear what the statistical power of the main reported effect (emotion faces) is.*
Wada et al. 2005b	*2/1/?*	*+/3*	*1 year*	*Same as Wada et al (2005), but including data on number of utterances. Silent PARO provokes significantly less utterances than normal PARO.*
Wada et al. 2005c	*2/1/14*	*+/5*	*20 minutes*	*Strong intervention and dubious interpretation of cortical neuron activation. Also only short term effect.*
Wada et al. 2005d	*2/1/8*	*+/2*	*17 months*	*Long term study but no new insights compared to the other work of the same group.*
Wada et al. 2006	*2/1/14*	*+/2*	*10 weeks*	
Wada & Shibata 2006	*2/1/11*	*+/1,3*	*1 month, 9h per day*	*Participants could play themselves with the robot without caregivers intervening. This is a clean study but does not have a good control group/situation. Social network increased in size and stress hormone indicated better immune system.*
Wada & Shibata 2007	*2/1/12*	*+/1,3*	*1 month, 9h per day*	*Participants could play themselves without caregivers intervening. This is a clean study trying to eliminate researcher intervention, however the control group/situation is not clear.*
Pearl				
Montemerlo et al. 2002	*2,4/4/6*	*+/5*	*5 days*	*An experiment with robot guidance.*
Pineau et al. 2003	*2,4/4/6*	*+/5*	*5 days*	*Experiment with elderly guidance using a robot. Same as Montmerlo et al (2002)*
Robocare				
Giuliani et al. 2005	*1/5,6/123*	*±/6*	*?*	*Evaluation of robot perception amongst elderly*

Design	Outcome measures		Task	
1. Comparative cohort	1. Health status	5. Other design criteria	1. Companion	5. Interface
2. Case studies	2. Mood	6. Remembering	2. Game/amusement	6. Physical help/ butler
3. Focus group	3. Communication	7. Acceptance rating	3. Information/coaching	
4. Narrative/ opinion	4. Loneliness		4. Guide/walking aid	

Appendix C – Overview of experiments

#	Description	Setup	Outcome
1	iCat social abilities using/testing UTAUT derived model & questionnaire (Chapter 3)	iCat robot WOZ in 2 conditions, location: eldercare institutions N=36 – more social 17, less social 19	R^2=,37 More Social ,28 Less social , 45 Model does not predict/explain differences
2	Annie social abilities using/testing UTAUT derived model & questionnaire (Chapter 3)	Screen agent WOZ Location: eldercare institutions N=33 – more social 17, less social 16	R^2=,59 More Social ,50 Less social , 65 Model does not predict/explain differences
3	iCat social abilities using new model & questionnaire justifying social constructs (Chapter 5)	iCat robot WOZ in 2 conditions, location: eldercare institutions N=40 – more social 20, less social 20	R^2=,70 More Social ,28 Less social ,45 Model does predict/explain differences
4	using new model & questionnaire justifying adaptivity construct (Chapter 6)	Robocare robot videos in four conditions – N=88 Neutral 22 Adaptable 21 adaptive + user control 23 adaptive - user control 22 Location: eldercare institutions and homes of older adults living independently	R^2=,68 1. neutral ,69 2. adaptable ,88 3. adaptive + user control ,69 4. adaptive - user control ,62 outcomes justify construct of perceived adaptivity third condition is most favorable
5	iCat public usage validating new model & questionnaire (Chapter 7)	'Autonomous' iCat robot with touch screen available for public use for 1 week N=30 Location: eldercare institutions	R^2=,63 Usage is predicted by Intention to Use
6	Steffie private usage validating new model & questionnaire (Chapter 7)	'Autonomous' screen agent on users computers for 10 days N=30 Location: homes of older adults living independently	R^2=,79 Usage is predicted by Intention to Use

Appendix D - UTAUT questionnaire

Performance expectancy
- *I would find the system useful in my job.*
- *Using the system enables me to accomplish tasks more quickly.*
- *Using the system increases my productivity.*
- *If I use the system, I will increase my chances of getting a raise.*

Effort Expectancy
- *My interaction with the system would be clear and understandable.*
- *It would be easy for me to become skillful at using the system.*
- *I would find the system easy to use.*
- *Learning to operate the system is easy for me.*

Attitude toward using technology
- *Using the system is a bad/good idea.*
- *The system makes work more interesting.*
- *Working with the system is fun.*
- *I like working with the system.*

Social influence
- *People who influence my behavior think that I should use the system.*
- *People who are important to me think that I should use the system.*
- *The senior management of this business has been helpful in the use of the system.*
- *In general, the organization has supported the use of the system.*

Facilitating conditions
- *I have the resources necessary to use the system.*
- *I have the knowledge necessary to use the system.*
- *The system is not compatible with other systems I use.*
- *A specific person (or group) is available for assistance with system difficulties.*

Self-efficacy
- *I could complete a job or task using the system.*
- *If there was no one around to tell me what to do as I go.*
- *If I could call someone for help if I got stuck.*
- *If I had a lot of time to complete the job for which the software was provided.*
- *If I had just the built-in help facility for assistance.*

Anxiety
- *I feel apprehensive about using the system.*
- *It scares me to think that I could lose a lot of information using the system by hitting the wrong key.*
- *I hesitate to use the system for fear of making mistakes I cannot correct.*
- *The system is somewhat intimidating to me.*

Behavioral Intention to Use the system
- *I intend to use the system in the next <n> months.*
- *I predict I would use the system in the next <n> months.*
- *I plan to use the system in the next <n> months.*

Appendix E – Almere model comprehensive toolkit

Introduction

The human-robot interaction community is multidisciplinary by nature and has members from social science to engineering backgrounds. In this 'comprehensive toolkit' we aim to provide human robot developers with a straightforward instrument to evaluate users' acceptance of assistive social robots they are designing or developing for elderly care environments. We will explain how we developed the measures for this analysis, provide do's and don'ts in designing the experiments, supply some suggestions on the application of the measures we have developed for this purpose and the analysis and interpretation of the data. As such we hope to engage developers in evaluating the acceptability of their own robot to inform the development process and improve the robot's design.

Model

The Almere model is a technology acceptance model: it can be used to predict and explain usage of a system by observing the influences on the Intention to Use this system. This Intention to Use predicts actual usage of the system – for some systems also Facilitating Conditions and Social Influence are predictive influences on usage.

The influences that are included in the model are used as variables and represented in a questionnaire by a group of questions or statements that can be replied to on a five or seven point or occasionally six, eight, nine or ten point Likert type scale. We used a five point scale in our experiments (totally agree – agree – neutral – do not agree – totally do not agree). To obtain the values for the variables scores can be attributed so that statistic processing is possible.

What makes this model usable for social robots and elderly users is that it includes specific influences representing social acceptance and the specific demands of elderly users.

As Figure F.1 shows, some constructs are directly determining Intention to Use while others are indirect determinators. A few are both.

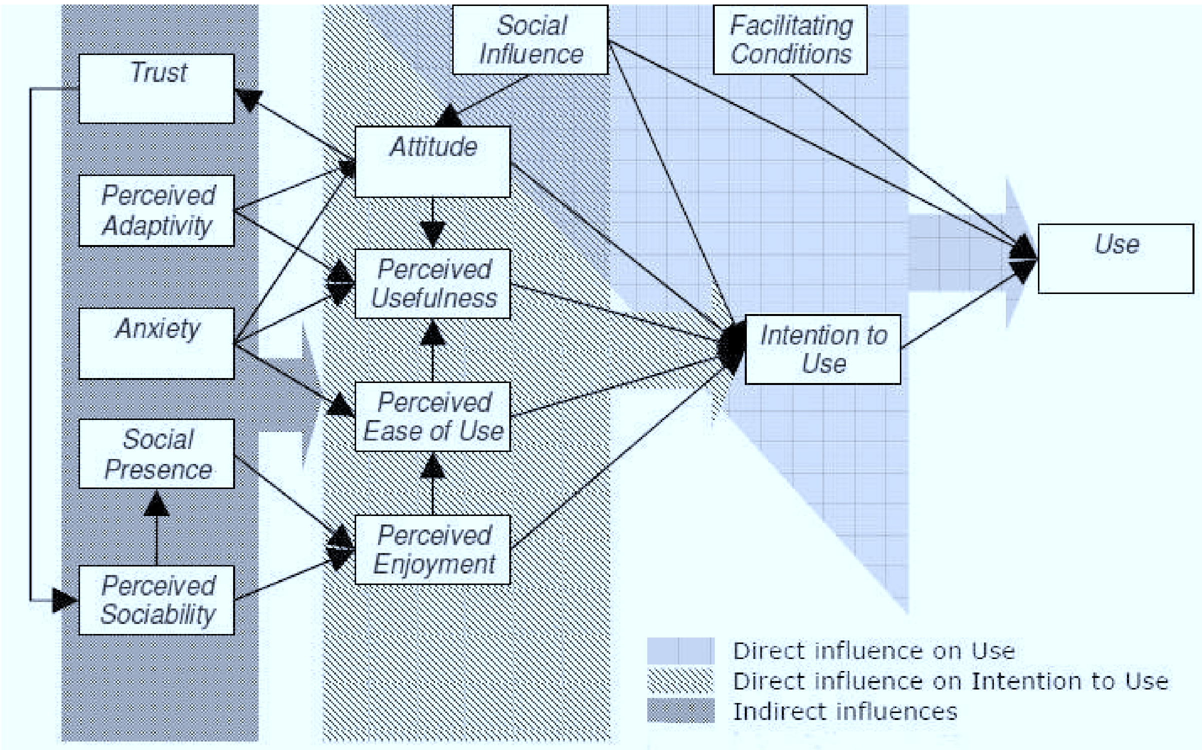


Figure F.1. Visualization of the Almere model

The model assumes the following construct interrelations:

1. *Use is determined by (a) Intention to Use, (b)Social Influence and (c)Facilitating Conditions.*
2. *Intention to Use is determined by (a) Perceived Usefulness, (b) Perceived Ease of Use, (c) Attitude, (d) Perceived Enjoyment and (e) Social Influence*
3. *Perceived Usefulness is determined by (a) Perceived Ease of Use (b) Perceived Adaptivity and (c)Anxiety*
4. *Perceived Ease of Use is determined by (a) Anxiety and (b) Perceived Enjoyment*
5. *Perceived Enjoyment is determined by (a) Social Presence and (b) Perceived Sociability*
6. *Perceived Sociability is determined by Trust*
7. *Social Presence is determined by Perceived Sociability*
8. *Trust is determined by Attitude*
9. *Attitude is determined by (a) Social Influence, (b) Perceived Adaptivity and (c) Anxiety*

These interrelations do not all apply for every system. On the contrary: acceptance of every specific system is characterized by the selection of the above interrelations that turn out to be significant.

Code	Construct	Definition	Determined by	Determining
ANX	*Anxiety*	*Evoking anxious or emotional reactions when using the system.*		*PU, PEOU, ATT*
ATT	*Attitude*	*Positive or negative feelings about the appliance of the technology.*	*ANX, PAD, SI*	*ITU, Trust*
FC	*Facilitating conditions*	*Objective factors in the environment that facilitate using the system.*		*Use*
ITU	*Intention to Use*	*The outspoken Intention to Use the system over a longer period in time.*	*ATT, PENJ, PEOU, PU, SI*	*Use*
PAD	*Perceived adaptivity*	*The perceived ability of the system to be adaptive to the changing needs of the user.*		*PU, ATT*
PENJ	*Perceived enjoyment*	*Feelings of joy or pleasure associated by the user with the use of the system.*	*SP, PS*	*ITU, PEOU*
PEOU	*Perceived Ease of Use*	*The degree to which the user believes that using the system would be free of effort*	*ANX, PENJ, PS*	*ITU, PU*
PS	*Perceived sociability*	*The perceived ability of the system to perform sociable behavior.*	*Trust*	*PENJ, SP*
PU	*Perceived Usefulness*	*The degree to which a person believes that using the system would enhance his or her daily activities*	*ANX, PAD, PEOU*	*ITU*
SI	*Social influence*	*The user's perception of how people who are important to him think about him using the system*		*ITU, Use, ATT*
SP	*Social presence*	*The experience of sensing a social entity when interacting with the system.*	*PS*	*PENJ*
Trust	*Trust*	*The belief that the system performs with personal integrity and reliability.*		*ITU, PS*
Use	*Use/Usage*	*The actual use of the system over a longer period in time*		*ITU, FC, SI*

Table F.1. Listing of constructs and assumed construct interrelations

Questionnaire

The questionnaire that has been developed alongside this model is listed in Table F.2. These statements are not to be presented in this order: they are randomly mixed and renumbered. Moreover, questions on age, education and (computer) experience can be added.

Procedures

If a system is tested to map all influences on its acceptance, the complete questionnaire can be used. A study can also focus on specific influences, which means only a part of the constructs is used.

The model or parts of it can also be adapted or even be used for other systems than assistive social robots. In that case it has to be re-validated: a significant relationship between Intention to Use and Usage has to be established.

ANX	*1. If I should use the robot, I would be afraid to make mistakes with it* *2. If I should use the robot, I would be afraid to break something* *3. I find the robot scary* *4. I find the robot intimidating*
ATT	*5. I think it's a good idea to use the robot* *6. The robot would make my life more interesting* *7. It's good to make use of the robot*
FC	*8. I have everything I need to make good use of the robot.* *9. I know enough of the robot to make good use of it.*
ITU	*10. I think I'll use the robot during the next few days* *11. I am certain to use the robot during the next few days* *12. I'm planning to use the robot during the next few days*
PAD	*13. I think the robot can be adaptive to what I need* *14. I think the robot will only do what I need at that particular moment* *15. I think the robot will help me when I consider it to be necessary*
PENJ	*16. I enjoy the robot talking to me* *17. I enjoy doing things with the robot* *18. I find the robot enjoyable* *19. I find the robot fascinating*
PEOU	*20. I find the robot easy to use* *21. I think I can use the robot without any help* *22. I think I can use the robot when there is someone around to help* *23. I think I can use the robot when I have a good manual.*
PS	*24. I consider the robot a pleasant conversational partner* *25. I find the robot pleasant to interact with* *26. I feel the robot understands me.* *27. I think the robot is nice*
PU	*28. I think the robot is useful to me* *29. It would be convenient for me to have the robot* *30. I think the robot can help me with many things*
SI	*31. I think the staff would like me using the robot.* *32. I think many people would like me having the robot.*
SP	*33. When interacting with the robot I felt like I'm talking to a real person* *34. It sometimes felt as if the robot was really looking at me* *35. I can imagine the robot to be a living creature* *36. I often think the robot is not a real person.* *37. Sometimes the robot seems to have real feelings*
Trust	*38. I would trust the robot if it gave me advice.* *39. I would follow the advice the robot gives me.*

Table F.2. Questionnaire

Processing results

Processing the results of the questionnaire usually includes the following procedure:

- Calculating the scores for each construct by averaging the scores on the items.

- Establishing Cronbach's Alpha for the items of each construct (Santos 1999). A reliable construct would have an alpha of at least .7 (Nunnaly and Bernstein 1978). If a construct consists of more than two statements, it is a good idea to see what the score would be if a question is omitted, especially if the alpha is not high enough. In that case, omitting a statement could be an option if solid arguments can be found.
- Analyzing basic descriptive statistics: minimum, maximum and mean scores, and standard deviation to get a first impression on the scores.
- Testing hypotheses with correlations (strictly explorative) or linear regression analysis (Montgomery et al. 2001; Tabachnick and Fidell 2001). A linear regression analysis would demand preferably at least 20 participants for each construct, but never less than 5. It can be performed separately for each hypothesis, which means a test would preferably include at least 120 participants and no less than 30. It measures how much each determining construct (being independent variables) is influencing a particular construct (being a dependent variable). Usually also an ANOVA table is generated with a regression analysis. This can be used to analyze the predictive value of the combined constructs within a hypothesis.
- The most profound way to analyze results would be to apply structural equation modeling. This could be used to establish alternative paths and the strength of construct interrelations. This would demand at least 15 to 20 cases (users) per construct though, and in this field it is often not possible to gather that many participants.
- Correlation scores can be used with any number of participants. They cannot be used to establish causal (determining) relationships, but as an indication that there is a relationship between two variables. Correlation analysis is especially useful if multiple tests need to be compared or if establishing determining relationships is not the subject of the study.
- Of course there can be additional statistics. When comparing different conditions or user for example, a t-test or Mann-Whitney u-test can be carried out.
- To test (parts of) the model, especially if any changes are made, also principal component analysis with rotation component matrix can be used to check if items that belong to a construct indeed 'load' on the same factor.

Assessing acceptance of assistive social robots by aging adults

Marcel Heerink