Social interactions in ambient intelligent environments

Social Interactions in Ambient Intelligent Environments

The work described in this thesis has been carried out at the Philips Research Laboratories Eindhoven as part of the Philips Research program.

An electronic copy of this thesis in PDF format is available from the TU/e library website (http://www.tue.nl/lib).

A catalogue record is available from the Eindhoven University of Technology Library
ISBN: 978-90-386-2403-7

Social Interactions in Ambient Intelligent Environments

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
rector magnificus, prof.dr.ir. C.J. van Duijn, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen
op woensdag 22 december 2010 om 16.00

door

Boris de Ruyter

geboren te Kortrijk, België

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr. E.H.L. Aarts
en
prof.dr. P. Markopoulos

Contents

1

Introduction

By nature humans develop from isolated and egocentric newborns into socially engaging human beings. The process of social development from a **psychological perspective** involves a separation from the self and the environment which is instrumental in the development of both personality and autonomy. This development is driven by social interest: a need to position ourselves in, and differentiate ourselves from, the social environment. In his *Individual Psychology*, Alfred Adler (Adler, 1938) discusses how social interest drives people to socially connect to society as a way to deal with an inferiority complex that characterizes the social development of human beings. Obtaining a position in society will deliver us status and recognition and thus help us in coping with this inferiority complex. Or, as Schutz (Schutz, 1958) argues in his *Interpersonal Needs Theory*: the drive to develop interpersonal relationships and communicate with others is based on the need to feel accepted, understood and valued.

Hence, one important aspect of our daily life is social interaction. However, this also introduces a serious tension: to satisfy our need for interpersonal relationships we risk being rejected by those from whom we seek being valued. In fact, the psychological cost of rejection increases as closeness [1] grows (Baumeister and Leary, 1995). Having intimate relationships will contribute to lower mortality rates, a lower risk for developing illness and less psychological symptoms (Prager, 1995). Given the tension between costs and benefits of engaging in social relations, it is obvious that managing social relationships requires some complex risk regulation mechanisms (Murray and Holmes, 2006).

From a **sociological perspective**, Charles Horton Cooley (Cooley, 1909) has explained the role of social interaction by means of the concept *looking glass self*. This concept describes a mechanism for developing our self-concept based on a comparison of our self perceptions and reactions we get from others during social interaction. The self-concept constitutes the knowledge and understanding we have of ourselves. Hence, the importance of social interaction during our development to establish our self-concept. From an evolutionary approach, social affiliation is assumed to be desirable for survival and reproductive purposes (Ainsworth, 1989). More specific, group membership is seen as a protective property that enhances human survival in, for example, situations of limited resources that need to be shared.

1.1 Social affiliation

Social affiliation is not only an essential aspect of our psychological and sociological development but it is also subjectively reported as a primary need (even above physical health) (Berscheid, 1985). Interestingly, the consequences and impact of social affiliation go far beyond psychological effects. For example, studies have found a relation between feelings of loneliness and increased unhealthy eating patterns (Hakwkley and Cacioppo, 2007), between social exclusion and reduced cognitive functioning (Campbell et al., 1982), between a non-satisfying partner relation and increased psychological and somatic health problems (DeLongis et al., 1988), between loneliness and a decrease in immunocompetence (Kiecolt Glaser et al., 1984), between not being part of a supportive social network and increased stress (Cohen and Wills, 1985). In general, based on an overwhelming body of literature, it is fair to conclude that deprivation from social connectedness can have a significant impact on both mental and physical well-being of human beings (Baumeister and Leary, 1995). While the studies reported in this thesis will not validate directly any of these impacts of social affiliation on mental and physical well-being, it is considered as an important starting

[1] the concept of closeness does not have a unified definition in psychological literature but is best described as a metaphor for the amount of sharing between persons (Schwartz, 1993)

point for investigating further the role technology could play in supporting social affiliation. The need to be socially affiliated or to belong is considered as one of the fundamental human needs (Maslow, 1968). It is assumed that satisfying the need to belong implies two necessary conditions: (1) there should be frequent and affective interactions with a few people and (2) these interactions must take place in the context of a caring relationship (Baumeister and Leary, 1995).

While both the objective and subjective importance of social affiliation has been clearly identified, it does not imply that all problems related to social affiliation have been resolved. For example, studies have reported that as much as 20% of the population continuously feels socially isolated (Steffick, 1985). Sociological studies (Weiss, 1973) have identified our changing lifestyle pattern as an important reason for the weakening of social affiliation. Within our current society, the strive for individualism has led to reduced social cohesion (Alperin, 2001). In a more recent discussion on the role of *social capital* [2] Helliwell et al. (2009) have argued for the individual and societal advantage of social affiliation. In his discussions on societal changes, Putnam argues for community engagement and social trust as measures of social capital (Putnam, 1993, 1995, 1996). Although the focus in these discussions is on institutional affiliations rather than individual relations, the importance of social affiliation for both the individual and society is reinforced.

1.2 Technology support for social interaction

While it is clear that technology cannot replace human-to-human interaction, it is worth investigating if Information and Communication Technologies (ICT) can support addressing interpersonal social needs where people are hindered in doing so by for example lifestyle patterns that limit frequent and affective interactions between people. Deploying ICT to enhance social interaction is not straightforward though. In his earlier analysis on the influence of technology on social capital, Putnam has pointed to technological changes (e.g. television) as responsible for a decline in social affiliation. Whereas the technologies Putnam was talking about lack the aspect of interactivity as we know it in today‘s ICT, a more recent study has investigated the impact of Internet use on social connectedness (Taube, 2004). In contrast to Putnam's operationalisation of social capital in terms of community engagement and social trust, this study has also considered the person-to-person social

[2] social capital is defined as networks together with shared norms, values and understandings that facilitate cooperation within or among groups

affiliation as an indicator of social capital. Interestingly this study cannot conclude that the introduction of ICT has the devastating effects on social capital as discussed by Putnam.

Following up on this potential role of ICT in social interactions, this thesis will present and discuss empirical studies on actually enhancing social affiliation by means of ICT. More specific, this thesis shall examine an emerging class of applications supporting sustained and endorsed social interaction between people. We shall refer to these applications as *social connectedness and awareness applications* meaning that their primary function is not to support private information exchange or person oriented communication but rather to (i) encourage social connection between individuals and (ii) to help them build awareness of each other.

These applications depend on continuous connectivity, for example by means of Internet technology, which has become possible with the introduction of broadband Internet technologies. To sketch the potential of these applications it is valuable to consider the broadband uptake in Europe. Despite large national differences, broadband penetration in Europe is making significant progress. The number of on-line households in Europe will grow from 103 million in 2008 to as much as 124 million[3] by the end of 2013 (Nuthall, 2007). This significant increase in the number of on-line households can have a significant effect on the introduction of social connectedness and awareness applications.

While there are still 11% of contemporary households in the developed world using dial-up connections, by 2013 this group will only represent 2% of the on-line households. Additionally, it has been found that the connectivity type (such as broadband) can have a significant effect on the types of services being picked up (e.g. a correlation of 86% between broadband adoption and on-line shopping). The importance of the shift from dial-up to broadband lies in the fact that such broadband access will now enable always-on services and allow for pervasive awareness applications to become integrated into our daily lives.

Another phenomenon is the uptake of broadband connectivity in Consumer Electronics. While there has been much discussion and speculation on the digital home of the future, it is still a debate if such vision will be real-

[3] that would be 63% of all households in Europe

ized through PC or Consumer Electronics devices. Today we observe a strong commitment of many Consumer Electronics manufacturers and chip-set companies to bring broadband connectivity to the Consumer Electronics market. According to Jackson (2008) the broadband Internet penetration in Consumer Electronics devices is driven by:

- ◇ *Easier setup and updates*. Today, updating firmware of Consumer Electronics devices requires shipping of DVDs to consumers. Connected Consumer Electronics devices would support easy setup and updating.
- ◇ *Access to content*. Bringing digital content to the consumer will be made truly possible by bringing connectivity to Consumer Electronics devices.

Major applications of broadband connectivity in Consumer Electronics devices and key players in bringing this to the market are: entertainment, communication, home automation, security and monitoring. As for today‘s install base, it is observed that in the US for example, one out of four on-line consumers has a home network. Again, an important fact when it comes to the assessment of potential for social connectedness and awareness services in the home.

Providing differentiating on-line experiences relies on aspects such as: (Bodine, 2007)

- ◇ *Exceeding people‘s basic needs*. Addressing people‘s needs is of course essential but in order to be differentiating with the social connectedness and awareness applications, it is important to offer more than just the basics of technical network connectivity.
- ◇ *Appeal to people‘s interests*. Although today‘s technologies would allow a multitude of functions and features, it is important that the proposition relates to people‘s actual interests. In social connectedness and awareness applications this means zooming in on specific interests such as presence information, availability information, etc.
- ◇ *Allow people to customize*. People want to be in control of their environment and of how they present their self to others. Providing a solution that is universal and static in both its behavior and interaction is reducing the attractiveness of the proposition. Adding customization options to social connectedness and awareness applications however does introduce new challenges in terms of for example the ease-of-use of these applications.

1.3 Thesis outline

This thesis is focussing on research into Ambient Intelligence. More specific, the thesis will focus on *infrastructures*, *methodology* and *applications* deployed in Ambient Intelligence research. By presenting work on those three aspects the thesis contributes to the scientific research field of Ambient Intelligence research. It should be emphasized that this does not imply that all Ambient Intelligence related research should be done within the presented infrastructures and with the presented methodology.

The applications of Ambient Intelligence presented in this thesis investigate the possible role of technology to support people in maintaining social connectedness. More specific, a series of empirical studies is reported that explore the potential for *Ambient Intelligent* environments to support and facilitate social affiliation through technology. By positioning the human need in the center of technology development, Ambient Intelligence (described in chapter 2) is an attractive approach to application development for social affiliation. The studies reported in this thesis have in large part been conducted in the ExperienceLab at Philips Research. These test facilities (described in chapter 4) are specifically suitable for experience research (described and illustrated with a case study in chapter 3) that requires controlled yet realistic testing environments. In a first series of experiments with *awareness systems* to mediate presence through silhouette representations of remote activities is investigated (see chapter 5) and mediating availability through presence information (see chapter 6). The results of these studies show that providing simple presence information can create feelings of social connectedness while supporting the function of availability judgements (i.e. using the presence information to estimate the availability of the remote person to engage into a conversation). Taking Ambient Intelligence (AmI) to the next level, the role of *social intelligence* (see the extended AmI model in chapter 2) is investigated in a controlled laboratory study and discussed in chapter 7. From this study it is learned that equipping technical systems with a level of social intelligence results in enriched and human-like interactions with the system.

1.4 Research hypothesis

While the work in this thesis is addressing the more general hypothesis that technology can support social affiliation, some more specific research hypothesis are defined and addressed in each chapter and tested by means of several controlled laboratory experiments. Some of these studies have a clear hypothesis on the expected effects of experimental manipulations, others have the character of **effect studies** in the sense that they investigate the effect of applications of technology for supporting social interactions in Ambient Intelligent environments.

1.4.1 Response bias due to the experimental environment

As will be discussed and argued in chapter 4, the Ambient Intelligence paradigm has introduced the need for testing facilities that simulate real-life environments to enable test participants in having so-called experiences (see chapter 3 for a discussion on experience research). These experiences are induced by the presence of Ambient Intelligence systems in the environment.

The hypothesis tested in this chapter is that **the experimental environment will not have a statistical significant influence on experimental data collected in this context when compared to a typical laboratory environment**. In other words, the experiment reported in chapter 4 assesses the potential response bias created by real-life like environments[4]. Involving 40 test participants, the hypothesis is tested in a *two-factorial repeated-mesaures study design* in which a typical usability study is conducted in the Philips Research ExperienceLab.

1.4.2 Mediating Social presence through silhouette awareness information

After having tested a research hypothesis on potential response bias due to the experimental environment, an experiment is conducted in the ExperienceLab to study the potential of technology in creating the feeling of social presence between remote[5] persons. More specific the hypothesis that **silhouette representations of a person in a remote context will create a feeling of social presence with that remote person** is tested. For this a *three-factorial within-subjects study design* of different visualization conditions was deployed. The experiment reported in chapter 5 involved 33 test participants and focussed on testing the research hypothesis for a group versus individual

[4]These environments are called Experience and Application Research Centers (EARC) and described as *advanced research infrastructures that support multi disciplinary teams of researchers and designers to study human - technology interaction and test new applications of technology*. The test facility within Philips Research that conforms to the EARC principle is the **ExperienceLab**

[5]that is: test participants not in the same physical environment

setting. As an addition to social presence the feeling of group membership is studied as a between subjects variable for the awareness information system based on silhouettes.

1.4.3 Mediating availability through silhouette awareness information

Taking the potential role of awareness systems to the next level, chapter 6 addresses the question if awareness information such as a silhouette visualization, can mediate the availability of a remote person to engage in communication. More specific, the hypothesis that **silhouettes of a remote person will provide sufficient awareness information to make a judgement about the remote person's availability** is tested. This hypothesis is particularly interesting for the challenge to stimulate social communication between remote persons. The research hypothesis is addressed in two experimental studies.

The first study, involving 44 test participants, followed a *two-factorial within-subject study design* of different visulization conditions to test two specific hypothesis:

1. **Silhouette visualizations are as good for making availability judgements as full video visualizations**
2. **Empathic skills of a test participant are related to making availability judgements**

The second experiment reported in chapter 6 continues the study into the potential of silhouette based awareness information for making availability judgements of a remote person to engage in a communication. In this study however the accuracy of the judgement was taken into account. The research hypothesis tested in this experiment is that **silhouette visualizations of awareness information are as good as full video visualizations for achieving empathic accuracy**[6]. This hypothesis is tested in an experiment involving 46 test participants.

1.4.4 Mediating social intelligence

The studies in the previous chapters have been focussing on the ability and effects for technology to create feelings of social presence between remote persons. In chapter 7 a laboratory experiment is conducted to assess the

[6] Empathic Accuracy describes a person's ability to accurately infer what another person is thinking or feeling

ability and effects of social intelligence in an interactive system. Although this study is more of an effect study, it addresses the situation where an interactive system in an Ambient Intelligent environment becomes an actor instead of a mediator. The context of this transition within the Ambient Intelligence model is discussed in chapter 2 and represents an important next step in the evolution of Ambient Intelligence.

To study the effect of social intelligence an experiment, involving 36 test participants, was setup in the Philips ExperienceLab. Trough a *one-factor between-subjects study design* the following research hypothesis were tested:

1. **test participants perceive the interactive system as having a level of social intelligence**
2. **an interactive system with social intelligence will have a positive effect on the test participant‘s *technology acceptance* and *satisfaction* when using this social intelligent system**

Publications for chapter 2

Aarts, E. and De Ruyter, B. (2009). New research perspectives on ambient intelligence. *Journal of Ambient Intelligence and Smart Environments*, 1(1):5–14

De Ruyter, B. and Pelgrim, E. (2007). Ambient assisted living research in carelab. *ACM Interactions*, 14(4):30–33

2

Ambient Intelligence

At the turn of the century, a distinguished group of researchers identified the potential devastating effects of rapid technological developments, as described by the generalized Moore's law [1], for the balanced relationship between humans and technology (Aarts et al., 2001). Whilst not ignoring the threats and risks of so called technology push (e.g. technology that would be introduced into people's lives without a meaningful application that could lead to a distantiation from technological innovation), the Ambient Intelligence (AmI) vision was introduced to emphasize the positive contribution these technologies could bring to people‘s daily lives. Within the AmI vision the serving of human needs is positioned centrally and intelligent applications of technology are seen as a means that enrich people‘s life. In coarse terms Ambient Intelligence refers to the *embedding of technologies into electronic environments that are sensitive and responsive to the presence of people*.

In Ambient Intelligence, the term **ambient** refers to technology being embedded in a large scale, in such a way that it becomes unobtrusively

[1] Moore's law states that transistor density in conventional integrated circuits has been doubling every 18 to 24 months since 1964. The *generalized Moore's law* refers to the assumption that this pattern of doubling density will also hold for future technological developments

integrated into everyday life and environments. Hence, the ambient characteristic of AmI has both a physical and social meaning. The challenge for ambient technologies is to become invisible while still providing meaningful functionality to end-users (Weiser, 1991). This challenge is not trivial to address and requires radically different approaches to human - system interaction [2].

Per definition, since AmI is following the approach of moving technology into the background, the user - system interaction paradigm will change (see Figure 2.1). Going through several phase, a transition from localized to distributed systems going together with a transition from technology oriented towards human oriented interaction is observed. When adding intelligence to interactive systems a part of the locus of control will shift from the user towards the system and this system will start taking initiatives based on the knowledge it has from the user and the interactive context (Bellotti and Edwards, 2001). The deployment of this intelligence requires the system to deal with the non-deterministic nature of people's ways of interacting with their environment. Initially aspects such as intelligibility and accountability are demanded from an intelligent interactive system. However, as the system is moving in its intelligence level from personalization towards anticipation, higher levels of intelligence (e.g. social intelligence) are needed.

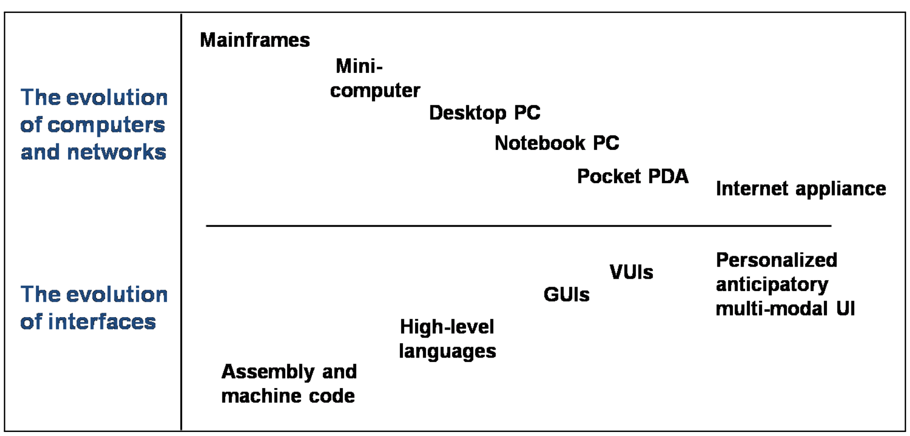


Figure 2.1: the changing interaction paradigm

[2]in chapter 7 a controlled laboratory study will be presented that has investigated to use of a social intelligent home dialogue system to overcome this invisibility challenge while supporting a fluent and acceptable flow of interaction between humans and the AmI environment

The term **intelligence** reflects the situation in which the digital surroundings exhibit specific forms of cognition, i.e. the environments should be able to recognize the people that inhabit them, personalize according to individual preferences, adapt themselves to the users, learn from their behavior and possibly act upon their behalf.

In AmI we distinguish between several levels of **system intelligence**: *context-aware*, *personalized*, *adaptive* and *anticipatory* system intelligence. Context-aware systems no longer ignore the environmental settings in which they operate. By having even an elementary level of awareness AmI systems can adjust their behavior to match the requirements of a context. While within the personalized level of intelligence, an AmI system can also adjust to a person‘s preferences, an adaptive system can take the dynamics of the context, personal and multi-user requirements into account. Ultimately, an anticipatory system will take initiative and make decisions on behalf of a human. Such a level of intelligence requires careful consideration and is prone for introducing undesirable consequences in the interaction between humans and AmI systems.

The user benefits of AmI are aimed at improving the quality of people‘s lives by creating the desired atmosphere and functionality via system intelligence and interconnected systems and services. Several futuristic scenarios have illustrated how technology can become supportive in people‘s daily lives (ISTAG, 2001). By mid 2006, a consortium of five European partners grouped under the name SWAMI focused further on some potential threats and risks of AmI scenarios. Although these scenarios are built around the same technological developments as those AmI was responding to, the scenarios are particular in their focus on addressing human needs. With these dark side scenarios, the SWAMI consortium emphasized that positioning human needs in the centre of technology development is not enough to ensure that the balance between humans and technology will be safeguarded (Wright et al., 2008). Although scenarios have been written and books have been published, the potential solution for this problem has not been provided other than suggestions for more technology (e.g. security related algorithms) development.

2.1 The extended AmI model

As technology and society are changing, the vision of AmI has also changed over the years. New requirements for the enabling technologies that relate to ethics, new methodologies for empirical research to better understand the context in which these applications will be positioned, a shift from system intelligence to social intelligence, are just some examples of challenges that call for a shift in Ambient Intelligence research (Aarts and De Ruyter, 2009). Elaborating on these issues, I discuss some important trends that influence not only the definition of Ambient Intelligence but also its research approach.

2.2 From entertaining to caring

Whereas Ambient Intelligence research has initially focused on user experiences in more entertainment/leisure/comfort oriented scenarios, there is a recent move towards the deployment of Ambient Intelligence technologies for Well-being and Care related application scenarios. Well being and Care applications cut across the domains of Lifestyle (e.g. persuasive fitness applications) and Healthcare (e.g. remote patient monitoring systems for chronic care patients). It should be clear that the development of applications related to our well-being and care will demand for some important shifts in the Ambient Intelligence research paradigm. More detailed presentation of possible Ambient Assisted Living (AAL) applications as well as future challenges for AmI research are discussed in De Ruyter and Pelgrim (2007).

2.3 From system intelligence to social intelligence

As AmI technologies are becoming more part of our daily life and aiming at promoting our well-being (e.g. through coaching and caring applications), there are increased expectations of AmI technologies to adapt and fit into social contexts. With this it is observed that the levels of system intelligence in the AmI paradigm require complementing with social intelligence. The earliest definition of Social Intelligence was coined by Thorndike (1920) and described as: the ability to understand and manage other people and to engage in adaptive social interactions. To adhere to and behave in a social intelligent manner is clearly a new challenge for AmI environments for which new destinations have been identified in the area of well-being and care.

When considering human-to-human social interactions, we see that there are several characteristics that make certain individuals stand out and more liked by others, or which convey an air of trustworthiness, competence and dependability (Ford et al., 1983; Keating, 1978; Sternberg and Smith, 1985). This list of social intelligent characteristics is large and includes attributes like *being nice and pleasant to interact with* and *being sensitive to other people's needs and desires*. In its broadest definitions social intelligence is **...a person's ability to get along with people in general, social technique or ease in society, knowledge of social matters, susceptibility to stimuli from other members of a group, as well as insight into the temporary moods of underlying personality traits of strangers** (Vernon, 1933). So the socially intelligent person has a better than average ability to judge other people's feelings, thoughts, attitudes and opinions, intentions, or the psychological traits that may determine their behavior. This judgment creates expectations on the observer's part about the likely behavior of the observed person. This in turn leads to adjustments of one's own behavior accordingly and appropriately. However, that appropriateness can only be judged when the social context is taken into account. In this sense, social intelligence is not merely something that goes on between two people in isolation, but contextual factors also come into play. The complexity and challenges for designing social intelligent systems are further discussed in Green and De Ruyter (2010).

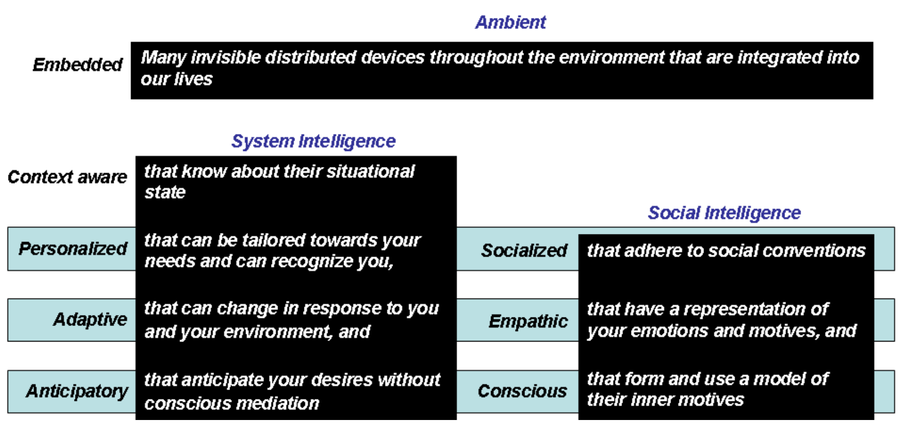


Figure 2.2: the extended AmI model

Social intelligence in AmI environments can take the form of a *socialized*, *empathic* or even *conscious* system (see figure 2.2).

2.3.1 Socialized

AmI environments that are socialized are compliant to social conventions. For example, in a sensing environment some form of system intelligence can be context-aware and thus know that a person is in a private situation. A personalized system would know that it is the user‘s preference not to be disturbed in such situation. An intelligent system that is socialized would use common sense knowledge to not allow disturbing the person in such a context. From this simple example it should also be clear that although we make a distinction between system and social intelligence at conceptual level, at implementation level both forms of intelligence need to come together.

2.3.2 Empathic

An empathic system is able to take into account the inner state of emotions and motives a person has and adapt to this state. Empathy is described as the intellectual or imaginative apprehension of another’s condition or state of mind without actually experiencing that person’s feelings (Hogan, 1969). For example, a form of system intelligence could infer that a person is getting frustrated while the social intelligent system with empathic capabilities would trigger the AmI environment to demonstrate understanding and helpful behavior towards that person.

2.3.3 Conscious

Ultimately, a conscious system would not only be aware of the inner state of the person but also about its own inner state. With such level of social intelligence, the conscious system could anticipate the effect a person is trying to get onto the system by reflecting on its own inner state. With this level of social intelligence it will be possible to develop rich and human like interactions in AmI environments. As with the anticipatory level of system intelligence, it will not be trivial to design AmI systems that can count on end-user acceptance. At this level of social intelligence one could be confronted with systems that develop for example their own ethical and moral standards.

2.4 Discussion

Ambient Intelligence is a vision on the development of technology applications with an emphasis on creating end-user experiences that highlight user benefits of new technology applications. Throughout the years of its existence, the vision had undergone several changes and the basic model of Ambient Intelligence has been extended with the notion of social intelligence in order to accommodate for the changing needs in society. It is clear that the Ambient Intelligence paradigm has a strong impact on the methodologies and instruments that are used for application driven research (e.g. designing and assessing end-user experiences). This impact is further discussed in the next chapter.

Publications for chapter 3

De Ruyter, B. (2003). User centred design. In Aarts, E. and Marzano, S., editors, *The New Everyday: Vision on Ambient Intelligence*. 010 Publishers, Rotterdam, The Netherlands

De Ruyter, B., Markopoulos, P., Aarts, E., and Ijsselsteijn, W. (2004). Engineering the user experience. In Weber, W., Rabaey, J., and Aarts, E., editors, *Ambient Intelligence*, pages 49–62. Springer-Verlag

De Ruyter, B., Van Loenen, E., and Teeven, V. (2007). User centered research in experiencelab. In Schiele, B., Dey, A., Gellersen, H., De Ruyter, B., Tscheligi, M., Wichert, R., Aarts, E., and Buchmann, A., editors, *Ambient Intelligence*, volume 4794 of *Lecture Notes in Computer Science*, pages 305–313. Springer

De Ruyter, B. and Aarts, E. (2010). Experience research: a methodology for developing human-centered interfaces. In Nakashima, H., Aghajan, H., and Augusto, J., editors, *Handbook of Ambient Intelligence and Smart Environments*. Springer

3

User Experience Research

The design of Ambient Intelligent environments differs markedly from the design of classical single device systems. AmI environments introduce new options for services and applications, by focusing on the desired functionality and the effect of that functionality on the user, rather than on the devices traditionally needed for each individual function. The fact that the technology will be integrated in these environments introduces the need for novel interaction concepts that allow the user to communicate with the AmI environment in a natural way. A cornerstone to Ambient Intelligence research is the design and creation of new user experiences[1] and the scientific investigation of their nature. When aiming at user experiences, empirical research for AmI environments has to take a step beyond the development of scenarios and the translation of use cases into system requirements. Although research methodologies such as User Centered Design (De Ruyter et al., 2007) remain important, much more radical departures are required. Rather than seeing interactive products as tools to perform certain tasks, user research must try to describe and nuance user experiences that help people enjoy, play, keep in touch with loved ones. Such experiences do not need to have a clear

[1] while the term *user experience* is still very much debated (Law et al., 2007), we describe a *user experience as the momentary emotional feeling elicited by interacting with applications of technology*

beginning and ending that is observable to the observer and perceived by the end-users (Abowd and Mynatt, 2000). They could span hours or days, and can sometimes result from infrequent, subtle interactions with technology that occupies a person's peripheral reach (Weiser, 1991). For this we propose an iterative empirical research cycle that consists of three phases: studies in context, laboratory and field (see figure 3.1).

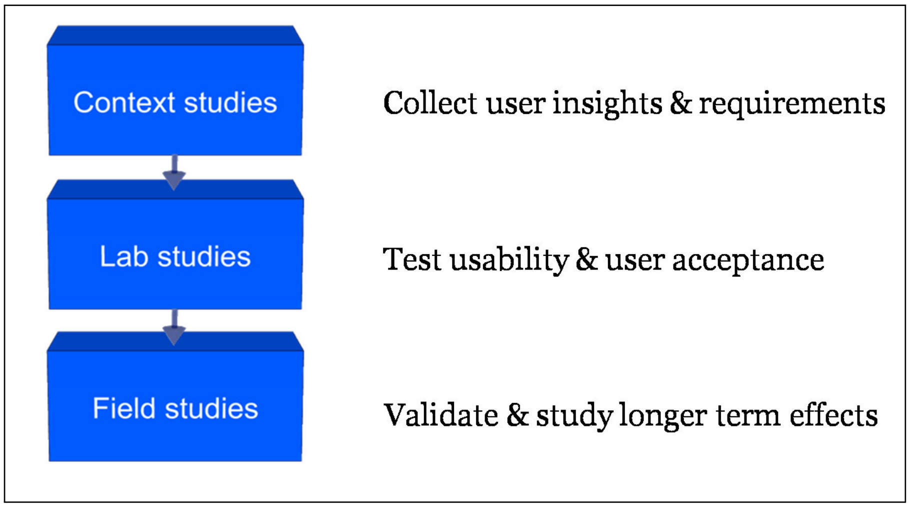


Figure 3.1: the User Experience research cycle

Although the presentation of the three phases might suggest a sequential approach, it should be noted that its implementation is iterative. From traditional User Centered Design cycles it is known that each phase in the research cycle will provide new insights but might also require the researcher to go back into one of the previous phases (De Ruyter, 2003).

The iterative Experience Research cycle consists thus of three phases: studies in (i) **context**, in (ii) the **laboratory** and in (iii) the **field**. Whereas the context studies focus on collecting initial user requirements without introducing any new technology applications, laboratory studies and field studies focus on the evaluation of new propositions in a controlled and real-life setting respectively. Whilst some studies reported very limited added value of conducting both laboratory and field studies (Kaikkonen et al., 2005), others have highlighted the added value of conducting field studies (McDonald et al., 2006; Kraut et al., 1996). Although Tory and Staub-French (2008) classify empirical studies in laboratory versus field settings as quantitative versus qualitative studies, we believe that such classification is an over simplification since both types of studies will highlight different aspects related to the user - system interaction and both types of studies allow for the collection of qualitative as well as quantitative data. The different phases are now further discussed.

3.1 Context Studies

In context studies the focus is on today‘s reality without introducing any new technology applications. By using ethnographic techniques (such as observations, in-situ interviews and diary studies), users are studied in their natural environment (Beyer and Holzblatt, 1998). Context studies can be seen as a way to understand the context in which future technology applications will be positioned. Although what is suggested as an approach is not participatory design, in which the end-user is not only the object of study but also the co-designer of new technology applications (Moss et al., 1927), such participatory design sessions can be conducted as a follow-up for applying the results of the context study in the design of new technology applications. As a technique for collecting contextual data, the context mapping approach (Short et al., 1976) has demonstrated to be very successful in gaining insight into the context of use for future technology applications. Also within the user system interaction research context, Rose et al. (1995) has demonstrated the added value of applying ethnographic techniques in context studies providing both qualitative and quantitative insights for improving interactive systems.

Gathering end-user needs by observing their behavior in context is a widely used method. Techniques such as experience sampling have been successfully deployed for studying for example end-user requirements of awareness systems in context (Khan et al., 2009, 2008). Such studies are aimed at understanding the contextual settings and requirements end-users might have for deploying awareness systems. Others have used more engaging techniques such as cultural probes[2] where test participants are invited to describe and express the associations they have with the home experience by means of for example drawings, picture taking, completing diaries and story telling (Eggen et al., 2003; Bernhaupt et al., 2007). Although ethnographic approaches have been widely used in the field of requirements elicitation for user - system interaction, the exploitation and integration of these contextual findings poses a challenge (Doherty et al., 2010; Postma et al., 2009).

[2] for a elaboration on the use of cultural probes see Gaver et al. (1999)

Once these studies have been completed, there is usually an overload of rich contextual data and the challenge is to abstract meaningful but not trivial insights without losing valuable information. It goes without saying that this is a difficult and cumbersome trajectory. All too often ethnographers complain that the step from the rich contextual data towards abstracted user insights is problematic due to information loss. Nevertheless, one needs to make such transition as the contextual data is too extensive to work with. One approach that has proven to be useful is to first compile some very high level insights from the contextual data and then to return to the rich contextual data to further understands these insights. This approach is very consistent with the approach to working with rich ethnographic data as suggested by Iqbal et al. (2005) and gives a focussed way of approaching and processing the rich contextual data.

For example (see figure 3.2), when studying the role of social networks in real-life settings (by for example using diary studies and observations), one can conclude, as a high level user need, that people need to receive appreciation from their social network. At first one might argue that such high level statement is rather obvious and that such insight was known even before conducting the contextual study. Or to quote Beyer & Holzblatt: *The complexity of work is overwhelming, so people oversimplify* (Beyer and Holzblatt, 1998). However, this is not where the analysis of the rich contextual data stops but rather where it begins: in the next step one will go back to the rich contextual data and explore the exact instances that have led to the generation of the high level statement. At that point one will be able to formulate user insights or concrete instances of user needs. In this example, one would go back into the rich data to understand how people experience and express appreciation through their social network. With this second step it will be possible to go beyond the obvious of the high level insights formulated during the first exploration of the rich contextual data.

Using this two step approach of abstracting and detailing it becomes possible to formulate valuable insights from the rich contextual data.

3.2 Lab Studies

System functionalities that generate true user experiences can be studied in a reliable way by exposing users to feasibility prototypes that provide proofs of concept. These are called experience prototypes, and they can be developed by means of a user-centered design approach that applies to both feasibility

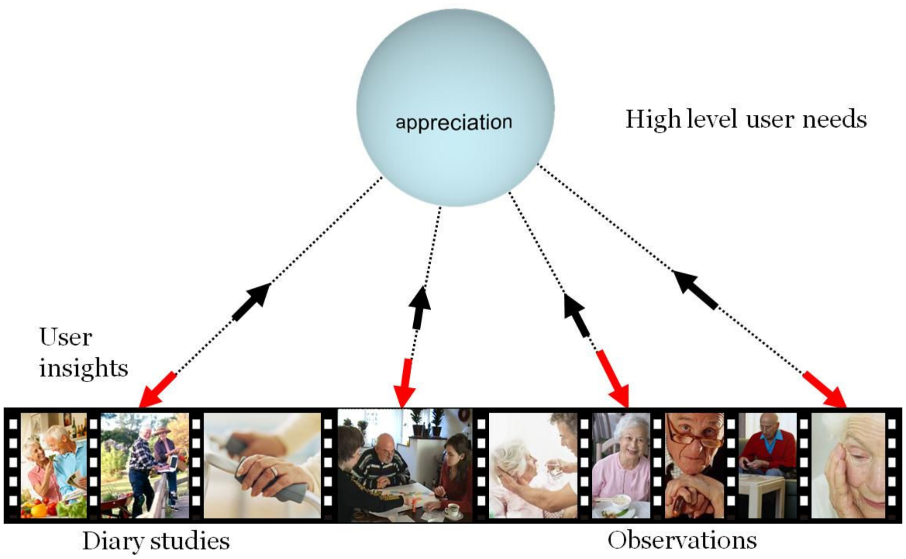


Figure 3.2: from rich context mapping data towards user insights

and usability studies in order to develop a mature interaction concept. For this laboratories are needed that contain infrastructures supporting fast prototyping of novel interaction concepts in environments that simulate realistic contexts of use. Moreover, these experience prototyping centers should also be equipped with an observation infrastructure that can capture and analyze the behavior of people who interact with the experience prototypes. Philips‘ ExperienceLab is an example of such an experience and application research facility (Aarts and Eggen, 2002). It combines the opportunity for both feasibility and usability research into user-centric innovation, leading to a better understanding of (latent) user needs and the technologies that really matter from a user perspective. The approach to conducting lab studies will be discussed in more detail in chapter 4.

3.3 Field Studies

Although studies in controlled laboratory settings can provide lots of valuable insights, the research findings are limited in terms of their ecological validity (Dix et al., 2004). With field studies the emphasis is on introducing new technology applications into realistic settings and studying the usage or behavioral change that might follow. In field studies end-users will be less enthusiastic about new technology applications and they will demand that these applications fit into their daily life by providing functionality that is meaningful to them. Additional, in field studies end-users will have the option to use these technology applications over longer periods of time.

There are no clear guidelines on the optimal duration of a field study. Some researchers (Neustaedter et al., 2007) have used pilot studies to estimate the needed duration of their field study in order to observe for example behavioral change. Others (Taylor, 2002) have used the term field studies to indicate both contextual studies (in which no new technology applications are introduced) and field studies in which technology applications have been introduced.

Although, in contrast to laboratory studies, there seems to be very little methodological guidance in conducting field studies, one can postulate the following guidelines:

- ◇ Field studies are often limited to the deployment of focused prototypes (rather than complete environments). This is both from a practical (i.e. installation & stability issues) and control perspective desirable.
- ◇ Field studies (although very much depending on the type of behavior that is being studied) will often spread from several days to weeks.
- ◇ Field studies will often be preceded by a period in which the users are not confronted with new technology applications. These periods serve as baseline for understanding the effect of the introduced technology applications. Often these periods will end with some data collection in the form of questionnaires. These questionnaires will be repeated at the end of the actual field test.
- ◇ Field studies rely for data collection mostly on logging data and in-situ interviews or questionnaires sampled over time. Although influenced by the complexity of the introduced technology application, it is found to be very useful to revisit and interview the user after an initial period of 3 days. After such initial period the end-users have experienced most of the application‘s functionality and have found the major obstacles in using the application. Interviews at the end of the field study will often not reveal these issues since end-users will have forgotten about them.

Rather than introducing abstract procedural recipes for deploying this framework the approach as advocated by Kuhn (1962) is adopted by introducing an example case study of how the framework is applied. The next sections will discuss this case.

3.4 Case study: Supporting social networks of elderly users

The particular case study focuses on supporting elderly users in maintaining contact with their social network. For this the concept of a community interactive television channel was implemented. Through this television channel the test participants can share information such as pictures and small messages with others tuned into this channel. Setup as an effect study, the focus was on observing how such an interactive channel would stimulate communication and create feelings of group cohesion.

3.4.1 Ageing society

Demographic trends are signaling the overwhelming need for ICT based consumer health and wellness applications. One of the high potential application areas is the elderly care domain: according to the World Health Organization, the worldwide proportion of people age 60 and over is growing faster than any other age group. By the year 2020 is it expected that 18% of the population will be over 65 years. Hence, there will also be a reduction in the available capacity of people that can provide care to seniors. This clearly points to an opportunity for technological solutions to support independent living for seniors. Additional, in chapter 2 we have seen that societal trends such as the ageing population are an important driver for applications in the Ambient Intelligence vision.

3.4.2 Elderly and social interactions

While well-being is a clearly multi-facet construct, it has been demonstrated that social networks are important for a senior‘s health (Antonucci, 2001; Seeman, 2000). Furthermore, research (Avlund et al., 2004) has indicated that social relations play an important role in preventing, to some degree, functional decline in aging. Social networks can be instrumental for both health and well-being of the individual (House and Kahn, 1985).

Whilst there are several types of social networks (i.e. social networks with family, friends or neighbors) there has been found a clear positive effect on a person’s well-being from social networks based on friends compared to social networks with only family (Fiori et al., 2006). Group membership provides several benefits to the individual such as for example deriving positive self-esteem (Tajfel, 1972). Social networks are the collection or maps of interpersonal ties that people of all ages maintain in varying contexts. Social networks have a structure (or topology) and interaction components. The whole of social networks of a person is called his or her social world. As we get older, there is a natural trend for social networks to shrink; elderly focus on the people that really matter to them, and let go other contacts (Baltes and Baltes, 1968). Also, the quality of social contacts is more important than the number of social contacts (Carstensen, 1995). Additionally, health problems cause a reduction of mobility which leads to less frequent contact moments.

3.4.3 Study Design

Following the proposed framework for User Experience Research, three research phases are followed:

1. **Context study**. Using an ethnographic approach, the social context of a selected sample of test participants is mapped. By creating maps of their social context a better understanding of their communication means and needs is gained. This leads to the definition of end-user requirements that can be addressed in a concept design phase.
2. **Lab study**. With the findings from the context study, a concept of a technology supported communication system is prototyped and tested in a controlled laboratory study. This study gives a better understanding of the usability and initial end-user acceptance of the concept and is setup as an effect study without any experimental manipulation. In other words, test participants are confronted with the concept‘s implementation and are asked to conduct a series of pre-defined tasks with that system. During this interaction information is collected on the usability of the system and, based on interviews, some first feedback on the end-user‘s acceptance for the proposed concept.
3. **Field study**. While a laboratory study provides valuable insights it does not represent a real-life situation. For this a field study will support observing how the concept is used in a daily life setting. Using interaction logging and post-field study interviews, information is collected on the effects of real-life usage of the system. Both results from the laboratory and field study are complementary and provide their own insights.

3.4.4 Test participants

The target group for this study is characterized by elderly (age 65 - 75), who are relatively healthy, have recently retired and who have an active social net-

Figure 3.3: interview

Figure 3.4: data analysis

work. The main focus was not identifying problematic situations (e.g. social isolation) but to better understand well functioning social networks of people who experience a major change in their life (i.e. retirement). Equipped with these insights research could conceptualize potential technology applications for elderly that are at risk of experiencing a reduced active social network without wanting to, and for which it is important to support them in maintaining and enjoying their social network. To further understand the structure and mechanisms of social networks in elderly based on friendship, we conducted a study consisting of three phases. These are now discussed.

3.4.5 Context study

After recruiting and introducing 11 elderly for the context mapping study, a probe package was provided to these participants. The purpose of this package is to sensitize (Sleeswijk Visser et al., 2005) the participants by making them conscious of their contact moments with others. More specific, the participants were requested to complete a poster that represents their social network. They were asked to work on the poster on a daily basis during a week, together with some assignments in a small workbook (Kang et al., 1996).

At the end of this first week, the participants were interviewed and the poster was discussed with them (see figure 3.3). Objectives of this interview included: get to know in what different social networks elderly participate, identify how these networks are structured and what motivations and feelings elderly have regarding maintaining and changing these social networks. During the interview it was agreed that participants would involve a companion (a person that plays an important role in their social network) for the second half of the contextmapping study, in which the research “zoomed in” on one social network per participant.

This activity was again setup as a collection of daily assignments for both the participant and his or her companion spread over one week. Afterwards, both would be interviewed to better understand their relationship and their common social network. During the interview they created a collage representing interactions and feelings of a specific activity they do together (e.g. drinking coffee, sporting together). At this occasion, there would also be a small interview on the use of currently available technologies (e.g. phone, email, postal mail) within the context of staying in touch with their social network.

In the next step of the contextmapping study, the findings of the contextmapping study are clustered and analyzed (see figure 3.4). The richness of this material requires making high level abstractions (step 1) that represent the high level needs of the participants. These are summarized as:

- ◇ There is a strong need for recognition from the actors in the social network; this gives the feeling of appreciation and satisfaction (see below)
- ◇ The elderly have a need for being and remaining independent, despite physical decline; they want to continue with the activities of their social world.
- ◇ The elderly experience general satisfaction about their social world, and don‘t need to extend it. They are bothered by the decrease in contact moments due to decreased mobility.
- ◇ After a major change in life, like retirement, divorce or relocation, social networks change. This causes a need for new social contacts, usually to replace‘ the ones which have been disappeared.

To support the further understanding of these abstractions, we focus on the underlying context mapping data that led to this abstraction (step 2). As such we have an efficient process for working with the extensive context mapping data. For example, the need for recognition was further detailed by specific behaviors and situations that were highlighted during the contextmapping study:

- ◇ **Being useful**. There is a strong need to make oneself useful; after retirement the mentioned group of elderly still feels responsible to add something to society to a certain degree.
- ◇ **Being included**. People want to be become member of an association or want to be invited at someone‘s birthday party. In other words, they feel the need to belong.

- **Confirmation of status**. People show what they are capable by comparing themselves with others with respect to condition, expertise or skills, or by challenging each other in a competitive way.
- **Mutual appreciation**. Recognition is expressed for other people‘s skills and expertise by giving each other compliments. Also the elderly show mutual interest in each other by exchanging news and ask for each other‘s situation. A lot of mutual help and support takes place as well, especially between neighbors: help each other with small jobs, cook for each other, and helping someone to his destination.

Although many of these findings are already confirmed in literature (Adler and Fagley, 2005), the present study provided more insights into how these fundamental human needs are addressed with today‘s solutions; it provides richness through the details and anecdotes the participants gave us. By first summarizing the highlight findings of the context study (step 1) and then going back to the rich context information related to these highlights (step 2) one obtaines a better focus on the rich context information. In a second level of the data analysis, it became clear how their social networks are formed, how the participants maintained and expanded these networks and what made these networks satisfying and rewarding.

3.4.6 Lab study

As a further research direction, the topic of social recognition and appreciation was explored. It is noted that the contextmapping data is rich enough for investigating many other aspects of elderly‘s social networks. Building on literature (Adler and Fagley, 2005), the following requirements for an application concept were put forward:

- Social networks consisting of few close friendships are preferred over large networks with less intimate friendships (Caldwell and Peplau, 1982)
- Satisfying social networks involve two criteria: (i) frequent and affectively pleasant interactions and (ii) these interactions must express an affective concern between the network‘s members (Baumeister and Leary, 1995; Romero et al., 2003)
- The social interaction should allow for small talk over trivial matters (Gerstel and Gross, 1982). Also, face-to-face contact was preferred over contact by telephone or email.

While these requirements are general and well documented in literature, the contextmapping study results supported understanding the context in which any technological solution would be introduced and additional requirements for the concept were established (e.g. the use of a system within a livingroom environment). After a further literature study and an exploration of existing technological solutions for supporting social networks, a concept creation phase was started. During this phase the elderly from the contextmapping study were asked twice to give feedback on the several developed ideas and concepts, after which these were further elaborated and refined.

The final concept developed in this case study is that of an interactive television channel in combination with a website, for deployment within one‘s specific social network (see figure 3.5), based on a common activity like Nordic walking or cooking. With this TV channel, the members of the social network can share short messages, pictures and slideshows that are meaningful to their network. In this way they can exchange short messages and share information they desire to share. The members of the community could express appreciation by attributing flowers to the postings (see figure 3.6). The actual posting is done using the website on a normal PC, while the consumption and sharing of the postings is limited to the TV channel. In this way, elderly with different levels of technology usage skills could participate.

The interactive channel would only allow a limited number of postings to be shared at the same time. As soon as new materials were introduced the older postings would be removed. The reason for this is that the channel is not meant to be a database; it aims to provide a quick, up-to-date overview of group content. It should enhance face-to-face contact, therefore the moment of use should be short.

The advantage of using the TV for this concept is that it is a familiar device for elderly. However, it also introduced a usability challenge for the concept‘s usability since all interaction have to be done using a remote control, in the most simple way. After consulting existing design guidelines (e.g. readability of text on TV displays) a first prototype of the concept was built and tested in the laboratory environment in terms of its usability and initial acceptance. The usability test was based on a comparison of actual usage logging and a pre-defined task model for the developed concept. More specific, test participants were asked to complete a pre-defined set of tasks while the effectiveness and efficiency of the interaction with the system was registered. Post experimental interviews revealed some of the participant‘s attitudes towards this concept. There were no major usability problems found while testing the

Figure 3.5: community channel

concept. Some minor usability issues with respect to the use of the remote control, the transitions between the messages and the readability of the posted messaged were identified. The concept was adapted and prepared as a robust prototype that could be deployed in a field setting.

3.4.7 Field study

The prototype (implemented as a set-top box that connects to any standard TV) was installed in six homes and was available for the users for a period of 8 days. After explaining the basic functionality and use of the system, the end-users could start using the prototype. The participants in this study were members of a small social network that organized frequent walking trips. The channel was used to share pictures and small messages related to their trips and gatherings.

Since the prototype was implemented as a networked application, detailed loggings were made of the daily use of the system. Throughout the field test, the participants were requested to complete a daily questionnaire measuring the perceived appreciation they experience from their social

Figure 3.6: appreciation rating

network (Adler and Fagley, 2005). The field test provided more insights into the use of the concept in real-life settings (see next paragraph). Additionally, it enabled researchers to investigate some specific mechanisms that enable and create strong social networks. In general, participants perceived it as the news channel' of their group. The storage of images was preferred over text, and they would have liked to store more than six items. Also, it was important for them to know the sender of an item. Based on the daily usage loggings we could observe that the system was intensively used by all community members. Analysis of the questionnaire data revealed that the system had a positive impact on the feelings of perceived appreciation. Confirmed by the post field test interview, participants indicated that they felt more connected to the community when using the concept on a daily basis.

The results of the field study highlighted some unexpected behaviors. For example, the rating with the flowers was not only used to express appreciation, but also as an indication that the item had been read. Also, content had been placed on the channel which was out of the scope of the group: its purpose was to show what other activities (in this case a weekend trip) this participant had.

Finally, one of the members of the social network was very negative towards the system due to its simplicity. After discussing the issue further it became clear that the person raising this issue was recognized in the social network as an expert when it comes to using ICT solutions such as the PC. However, due to the concept's simplicity any member of the social network could post and retrieve messages and pictures. As a consequence, the specific user felt that his status as expert was compromised since his support was no longer needed.

3.5 Discussion

The presented case study followed a three step approach consisting of a *context*, *laboratory* and *field* study. Each phase provides complementary insights needed to develop a concept of a system to support social networks amongst seniors. Although the contextmapping study is qualitative by nature, the rich and multi layered information, full of anecdotes and details, delivered by this study provided a basis for concept development. The concept was meant to support the expression of appreciation, as a means to meet the elderly‘s need for recognition. This concept was next implemented as an interactive television service and evaluated in a controlled laboratory setting. Based on the results of the laboratory test, the concept was improved and prepared for a small field trial. In the field trial, seniors would have the option to use the concept over a longer period of time (in this case 8 days instead of a one hour session in the laboratory environment). While this time span is most likely not enough for the prototype to fully embed into their routines (or to even change their routines), it does provide more insight into the use and appreciation of such concept.

The findings of this study support making following conclusions:

- ◇ The **context mapping study** provides information about more generic but technology independent needs and indicates that there is a strong need for recognition from the actors in the social network; this gives the feeling of appreciation and satisfaction
- ◇ A **laboratory study** provides information regarding the actual usage of applications of technology and indicates that the use of an interactive television has several advantages. First, the threshold for using this technology is low and seniors feel comfortable using this device. Second, the use of a television to broadcast a community channel provides the seniors with an acceptance that all posted information is shared with anybody having access to this broadcast channel.

◇ The **field study** provides information on the context of use and potential social consequences related to the usage of the applications of technology and confirmed the value of the proposed concept. More specifically, both the usage analysis of data collected during the field study and the post field study interviews revealed that the interactive community channel was intensively used and led to stronger feelings of connectedness.

In this chapter a *conceptual structure* or *framework* for conducting Experience Research is proposed. Following principles form the USER - CENTRED DESIGN (UCD) approach (De Ruyter, 2003) the framework proposes an iterative cycly between context - laboratory - field studies. Given that the methodological approaches differ per research phase in the proposed framework, more procedural recipes are available in the literature in the field of sociology, anthropology and psychology. Table 3.1 presents some of the differences between the phases.

	Context	**Laboratory**	**Field**
focus on	current reality	use of artefact	use of artefact in context
amount of data	more than expected	expected	expected plus anecdotes
type of data	pictures, drawings, stories, ...	numbers	numbers and anecdotes
scope of data	broad and diverse	focussed and specific	enriched but specific
methodological challenge	accuracy and precision	validity and reliability	validity and reliability
related discipline	sociology and anthropology	psychology	applied psychology
approach	elicitation	experimentation	quasi experimentation

Table 3.1: characteristics of the CONTEXT - LABORATORY - FIELD phase in the experience research framework

As argued by Dourish (2007) one should consider a methodology in a broader sense than a technique to be used for data collection. Beyond providing a technique or method for data collection, a methodology implies also taking into account the scope and limitations of knowledge that can be developed by adopting a methodology. Hence, in that line of reasoning, it is here argued that adopting the methodological framework of studies in context, laboratory and field environments results in complementary findings with each their scope and limitations. The risk of applying the proposed framework without considering and accepting this broader interpretation of methodology is that the researcher ends-up with an eclectic approach that does not provide the optimal level of scientific insights.

Demonstrated with the case study in this chapter it becomes clear that of the phases in the context - laboratory - field research cycle provides complementary insights that can be built upon in the consequent phase. Acknowledging that more structured methods and techniques are needed to strengthen the intertwining of these phase, the contribuition of the reported work is the formulation and illustration of the methodological framework for conducting user experience research in an Ambient Intelligent application context.

Publications for chapter 4

De Ruyter, B., Van Loenen, E., and Teeven, V. (2007). User centered research in experiencelab. In Schiele, B., Dey, A., Gellersen, H., De Ruyter, B., Tscheligi, M., Wichert, R., Aarts, E., and Buchmann, A., editors, *Ambient Intelligence*, volume 4794 of *Lecture Notes in Computer Science*, pages 305–313. Springer

De Ruyter, B., Van Geel, R., and Markopoulos, P. (2009). Measuring the response bias induced by an experience and application research center. 5859:235–244

De Ruyter, B. and Pelgrim, E. (2007). Ambient assisted living research in carelab. *ACM Interactions*, 14(4):30–33

4

ExperienceLab

After introducing AmI and some of its most recent changes in chapter 2 and the novel way of working introduced in chapter 3, we now introduce the instrument to implement this way of working: the **ExperienceLab**. From the discussion in chapter 2 it follows that for pursuing end-user experiences by means of applications of technology, both the research methodology and instrumentarium should be tailored towards this ambitious goal. Although it was stated in chapter 3 that the user experience is elicited by the applications of technology, the influence of the contextual settings of these applications should not be underestimated. While there are no cited scientific publications reporting on quantified effects of contextual settings on the end-user experience, it is highly acceptable to assume such effects. Ethnographic research methodologies such as those described in chapter 3 reflect the importance of contextual settings.

To demonstrate the intuitiveness of this assumption one could consider the example of a research study into a person's *wake-up experience.* Studying such experience does not only require a person to actually undergo this situation, it also requires a specific context. Instructing people during a study to imagine a wake-up experience in a laboratory setting will have limited scientific value.

The outline of this chapter is as follows: at first the ExperienceLab will be briefly introduced. Next, this chapter will report on a controlled experiment to test the research hypothesis that the experimental environment will not have a statistical significant influence on experimental data collected in this context when compared to a typical laboratory environment.

4.1 ExperienceLab

Environments that support the simulation and investigation of futuristic contexts have been arround for some time now. Many of these environments have been established from a technological perspective[1] (Brumitt et al., 2000). The **PlaceLab** at MIT (Intille et al., 2005) is a condominium in Cambridge, Massachusetts, that offers a controlled home environment (in the form of an appartment) for studying how people interact with new technologies. The main form of data collection is through sensors installed throughout the appartment. Besides studying interaction with technology, the PlaceLab also investigates architectural aspects of future living environments. The **Aware Home** (Kidd et al., 1999; Kientz et al., 2008) at Georgia Tech started as a research environment in the form of an authentic home environment driven by a research agenda. Although also investigating the role of technology in future living environments, the research at the Aware Home has a clear application focus and has been instrumental for research into application domains such as those described in this thesis. Focussing on both the ambient intelligent environments and the related building systems, the Fraunhofer **In-Haus** (Meyer and Vom Bögel, 2008) in Duisburg offers a testing environment for home, office, hotel and residential living contexts. In contrast to the MIT LivingLab and the Georgia Tech Aware home, the InHaus operates mainly in cooperation projects with business partners having market innovations as scope.

As an instrument for AmI research we will now discuss the Philips ExperienceLab. The ExperienceLab currently consists of a HomeLab, ShopLab and CareLab and acts as an innovation center for the development of novel consumer products and services, and therefore makes a substantial contribution to the implementation of the Philips strategy in the domain of Lifestyle and Well-being. While the ExperienceLab shows great communality with facilities such as PlaceLab, Aware Home or InHaus, there are some clear differences. First, in contrast to the PlaceLab, the ExperienceLab uses cam-

[1] for a discussion on such facilities and example applications see Li Jiang et al. (2004)

eras for complementing data collection through sensors. Second, the focus on applications and their embedding into a realistic environment is of greater importance than the mentioned facilities. Third, in contrast to InHaus, the focus is on applications of technologies that are further away from market introduction and for which additional technology development might be needed.

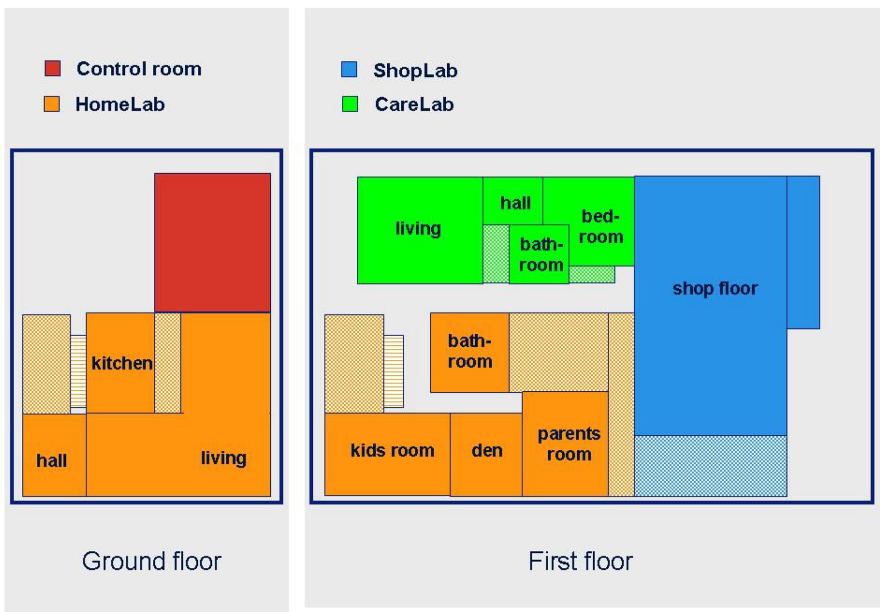


Figure 4.1: the ExperienceLab floor plan

4.1.1 HomeLab: the home environment

The HomeLab is built as a two-stock house with a living, a kitchen, two bedrooms, a bathroom and a study. At a first glance, the home does not show anything special, but a closer look reveals the black domes at the ceilings that are hiding cameras and microphones. Equipped with 34 cameras throughout the home, HomeLab provides behavioral researchers an instrument for studying human behavior inside HomeLab [2]. Adjacent to the Home there is an observation room. When HomeLab was opened in 2001, one way

[2] Important in such an environment is to point out to participants where the cameras are that will be used for recording (after they have given their informed consent). By doing so, participants will not be puzzled by the presence of the cameras and ignore them after a few minutes already.

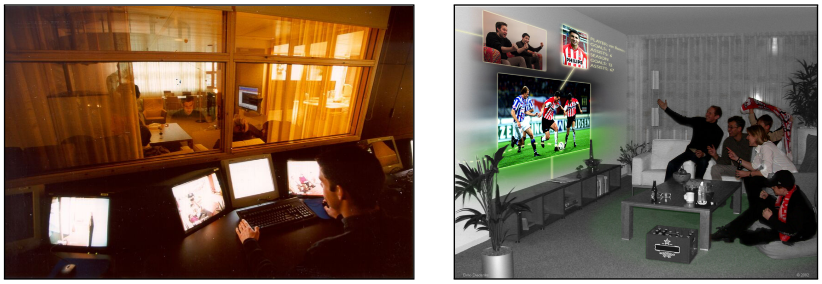


Figure 4.2: on the left the observation system and on the right the interior of HomeLab

mirrors were placed between the living and observation room. The idea was to have a direct view into the living. However, time taught that observers preferred the camera images. The different viewing angels and possibility to zoom into details were reasons to abandon the mirrors. The observation room is equipped with four observation stations. Each station has a large high resolution flat screen showing a collection of six different images from the cameras in the house. The observer is free to choose which cameras he wants to use and what the pan, tilt and zoom position of every individual camera has to be. And he can route two of the roughly 30 available microphones to the left and right channel of his headphones. Each observer has an application running to feed the behavioral data to the storage system, synchronized to the video data. In the early days of HomeLab this setup was used in real-time situations: while observing test participant‘s behavior inside the HomeLab observers would annotate the video signals with event data. Although the technical characteristics of the observation system allows for real-time coding by simultaneous researchers, the observers had a hard time in directly following the progress of the experiment. Nowadays it is more common to first have the video data stored on the capture stations and do the data annotation afterwards as realtime annotation is very work intensive with the risk of missing events due to time pressure. Also events and sensor data are time-stamped and appended to the video data. This way of working is much more efficient and a single observer can collect all the relevant data for further analysis once the data collection session has been completed.

Broadband Internet facilities enable various ways to connect parts of the HomeLab infrastructure to the Philips High Tech Campus network or even to the outside world. A wireless Local-Area Network (LAN) offers the possibil-

Figure 4.3: on the left test participants interacting with the shop window and on the right interacting within the ShopLab

ity to connect people in HomeLab without running cables. However, if cables are required, double floors and double ceilings provide nice hiding places. Corridors, adjacent to the rooms in HomeLab, accommodate the equipment that researchers and developers need to realize and control their systems and to process and render audio and video signals for the large flat screens in HomeLab. Light control systems (through a Local Operated Network (LON) protocol) can be accessed by the researchers and offer their prototypes the possibility to affect the light settings in the rooms.

4.1.2 ShopLab: the retail environment

The ShopLab research program builds on the insight, that shopping itself has become an important leisure activity for many people, and that flexible atmospheres are needed to enhance shopping experiences. On the other hand many retail chains want to maintain a clear house style for branding reasons. This introduces the challenge of combining these two major aspects. One approach to this, studied in ShopLab, is that one atmosphere design will be sent to all stores and slightly adapted there to meet local conditions. With the introduction of solid state (LED) lighting, a wide range of new options to create such atmospheres using color and dynamic effects is becoming available. However, tuning these atmospheres requires controlling several hundred lamp settings, introducing a complex overall control challenge. Another approach studied to enhance the shopping experience is the introduction of interactivity, in the form of interactive shop windows, interactive signage and reactive spots. Adaptation of these shop atmospheres also requires input from smart environments that detect people‘s presence and product interests while they are in or near a shop.

The ShopLab is used extensively to perform user studies, both with retail-

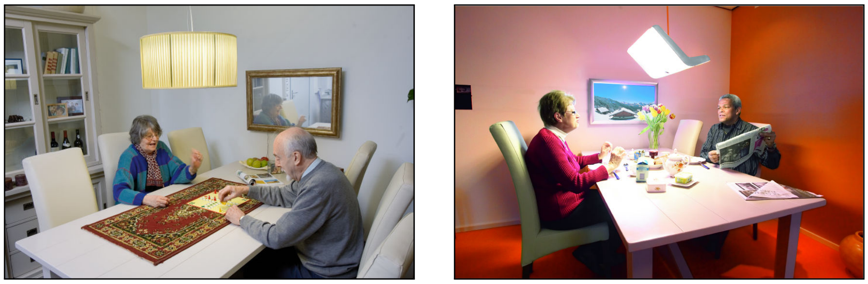


Figure 4.4: test participants exposed to activity- and circadian rhythm influencing lighting in CareLab

ers and with end-users (shoppers). By involving these users in all phases of the design process, including the evaluation of the experience prototypes, important insights in the actual experiences of users are obtained early on in the development process. For a discussion on the design of ambient experiences in a retail environment see Van Doorn (2009).

4.1.3 CareLab: the assisted living environment

This CareLab resembles a one-bedroom apartment for seniors and is equipped with a rich sensor network to study the contextual settings in which people will use the health and wellness applications (De Ruyter and Pelgrim, 2007). The sensor information is processed and combined to extract higher-order behavioral patterns that can be related to activities and states, such as the presence of people, the state of the home infrastructure, etc. With the CareLab it is possible to explore at an early stage the user's acceptance for these solutions and to assess the interactive and functional qualities of these solutions before deploying these into field settings. Results are used to improve applications of innovative technologies, to eliminate imperfections and to explore new applications.

4.1.4 Using the Experience Labs

When setting up an experiment in ExperienceLab, the researcher designs a coding scheme for the observation session. A coding scheme lists all prototypical behaviors that are expected to be observable during the session (Bakeman, 2000). These behaviors should be structured as an orthogonal classification system: during the classification of behavior it should not be possible to classify one behavior in more than one category. The observers mark the occurrence of these behaviors during the ExperienceLab session (or offline coding for recorded sessions) by means of pressing keys on a keyboard. Additional, more and more behavioral events are registered automatically through the use of sensors.

Once the coding scheme is developed it is saved into the scoring system. Very similar to questionnaires, the coding schemes are standardized and reused. For example, coding schemes for problem solving behavior or user-system interaction in voice controlled environments have been developed and reused over several experimental sessions (for an example of a reusable coding scheme see Bakeman and Casey (1995); Bakeman and Gottman (1997)). If applicable, a detailed user profile for test participants is established in collaboration with a consumer marketing intelligence department. The user profile is then provided to a recruitment agency for test participant recruitment. As a rule these participants are externally recruited to not have any affiliation with Philips as such affiliation could influence test results. Depending on the focus of the research question, a test methodology is designed.

Data analysis can consist of a simple frequency analysis up to a sophisticated data mining analysis for finding hidden patterns in the data set (Magnusson, 1996). For this ExperienceLab is equipped with a software tool capable of detecting repeated patterns that are hidden to observers and very hard or impossible to detect with other available methods such as for example cluster analysis. It is particularly suitable for analyzing behavioral data. This tool is able to detect patterns that are obscured by other events, and finds patterns that no form of frequency count, lag sequential or time series analysis can identify. As such, it is an effective way to detect patterns in user-system interaction and to identify the precursors or consequences of specific behavioral events. This tool has been used extensively in studies of human communication, spoken dialogues, gestures, protocol analysis and human-computer interaction (De Ruyter and Van Schijndel, 2004).

4.1.5 Lessons learned on the use of Experience Labs

Since the opening of ExperienceLab in 2001, there are several lessons learned with respect to the use of the ExperienceLab.

- ◇ Real-time observation coding is less important, off-line coding is preferred. When equipping the ExperienceLab with observation tools, it was assumed that researchers would code observations in real-time. However, over the years we have learned that off-line coding (i.e. after the experiment) is preferred due to the high time pressure involved in real-time coding. This has a consequence for the way data is collected and made available for scoring since now researchers need the option to export the observational data from the ExperienceLab system.

- ◇ Developing good coding schemes is as much effort as developing a questionnaire. Coding schemes provide an extensive classification of potential observable behavior. This coding scheme is used to code the recorded behavior. Developing good coding schemes takes time and reuse of these coding schemes (like for questionnaires) is desired.
- ◇ New methods and instruments to measure the subjective user experience in an more holistic way are needed. Although the user experience is by nature subjective, there is a need to capture and analyze the broader context of user experiences. Examples of new methods include the collection of psychophysiological data from the test participants.
- ◇ ExperienceLab is a catalyst for improving technology transfer into the business. Traditionally, research results are communicated through scientific publications and presentations. Over the years ExperienceLab has proved to be a very effective communication tool within a large corporate environment. Although the original goal of the ExperienceLab was to support end-user and feasibility research, there is a need to reserve capacity for demonstration and dissemination events.
- ◇ Having a support team is essential when operating an ExperienceLab. Since the opening of the ExperienceLab there has been a permanent software engineering team available for technology integration and maintenance of the infrastructure. Similarly, there is a need for a team of behavioral scientists to guide the empirical research in ExperienceLab.

In recent years we have seen the rise of Experience and Application Research Centers (ISTAG, 2004). An Experience and Application Research Centers (EARC) is best described as *an advanced research infrastructure that supports multi disciplinary teams of researchers and designers to study human technology interaction and test new applications of technology*. A differentiating aspect of these infrastructures is their high level of realism to simulate real life settings and the sophistication of technological embedding. In contrast to traditional research laboratories, these environments aim at providing a naturalistic context for end-users to engage in interactions with applications of advanced technologies. In traditional research laboratories the emphasis is on the standardization of the research context and the removal of any potential additional influencing factors that could stem from the testing environment.

4.2 Response bias in Experience Labs

While these EARC have been successfully deployed in evaluating the user-system interaction with applications of advanced technologies (Russel and Cousins, 2004), the response bias elicited by these infrastructures as assessment tools on the collected data, has not been studied. Yet, in other research areas interested in assessing the user experience, it is known that the contextual settings can be of influence on the collected data (Meiselman, 1996; Petit and Sieffermann, 2007; Bell et al., 1994). Given that the object of research in these environments is often the user's (subjective) experience and many assessments are based on self-reporting, it is important to investigate the effect a testing environment can have on these assessments.

One concern that is raised with a testing environment such as the ExperienceLab is thus its potential influence on the collected data. This concern is strengthened with the observation that end-users inside such environments have high expectations of finding advanced technological applications. In one anecdotal situation a test participant would inspect all furniture inside the ExperienceLab in order to find the sensors as this user was convinced that such an environment would be packed with technology. To evaluate this potential threat for empirical research, a controlled study was conducted in which a typical usability study was replicated in both a standard and low tech environment (see Figure 4.5) and the HomeLab (see Figure 4.6).

Response bias can be described as *any systematic tendency of a respondent to manifest particular response behavior for an extraneous reason which is not part of the experimental manipulation*. In the past, various researchers (Donovan and Rossiter, 1982; Mehrabian and Russell, 1974; Kotler, 1974) demonstrated the influence of specific facets of the environment on consumer behavior. To evaluate potential response bias, a controlled experiment was conducted in an EARC environment. This experiment involved the replication of a traditional user test (of an experimental system for video editing through the use of an interactive television set) in both the Philips HomeLab and a traditional laboratory environment. While the application to be tested was exactly the same in both environments, the EARC environment presented a better aesthetic integration of the system into the environment and provided an overall more natural situation. The experiment and its findings are discussed below.

Figure 4.5: UsabilityLab

Figure 4.6: HomeLab

4.2.1 Experiment

Method

Part of the experiment was conducted in the Philips HomeLab as described in section 4.1. A total of 40 participants, who had not visited the HomeLab before participating in this study, were recruited. The experiment was designed as a within subject design consisting of two experimental sessions separated by a time interval of one week. It was suggested to participants that there would be an improvement of the system's usability between the sessions (see Table 4.1) based on their input during the first session. In reality however, only minor changes that would not influence the system's usability (e.g. minor changes in the user interface colors) were made to the system in order to be able to compare the findings from both sessions. By implementing minor changes, that are enough for test participants to note a modification to the system, it was assumed that users would not strive for giving consistent answers to their earlier ones but would rather evaluate the system from scratch. Any changes in reported experience between the two sessions can then be attributed to potential response bias. It was assumed that a one week interval would be enough to not have any memory effects from the first session. In both sessions we conducted a typical usability (i.e. assessing the test participant's usage and appreciation of an interactive system) test of the same experimental system.

Procedure

Setup as a typical usability test, the ease of use of a video editing system through an interactive television set was assessed in both the EARC and in a traditional laboratory environment. The laboratory environment is less at-

	Session 1	Session 2
Group 1	EARC	EARC
Group 2	EARC	Laboratory
Group 3	Laboratory	EARC
Group 4	Laboratory	Laboratory

Table 4.1: Experimental design

tractive and less realistic compared to the EARC environment. The most important difference between both environments is the aesthetic embedding and presentation of the experimental system as part of a home like environment. Other properties of the system (e.g. the interaction and display devices), the environmental settings (e.g. lighting conditions) and the experimental procedure would be the same for both environments. A set of questionnaires was preceded by a short introductory text that differed depending on the experimental setting the participant was about to enter. In the introduction participants were told they were taking part in a research on the usability of a recently developed video editing system. Consequently, participants were given some information on the system to make it possible that certain expectations could be evoked. Next, the participants were told the usability test had been divided in two sessions with one week in between in which the video editing system would be adjusted on the basis of first session's results. During the second session the participants were asked to evaluate the system again.

Apparatus

The experimental system deployed in the usability test is a TV-based video editing system that is able to automatically convert a home video into an edited version, which is essentially a summary' of the raw footage. By means of a remote control, the user can edit and modify this automatically created summary by, during the viewing of the summary, pausing the desired shot and subsequently selecting one of the editing functions (e.g. adding music, adding effects).

Measures

During the usability test, participants were requested to complete a given set of tasks with this system. Several usability and contextual measures were collected during the experiment. The psychometric instruments used for data collection are discussed briefly.

- *Brief Mood Introspection Scale (BMIS)*
  In order to measure participant's mood (since mood could have an influence on the appreciation of the environment and experimental system), the Brief Mood Introspection Scale (BMIS) by Mayer and Gaschke (1988) was administered[3]. The BMIS consists of 16 adjectives which are based on eight mood states: (1) happy, (2) loving, (3) calm, (4) energetic, (5) fearful/anxious, (6) angry, (7) tired and (8) sad. Participants were asked to complete this questionnaire before starting the actual experiment.
- *Pleasure, Dominance and Arousal scale (PDA))*
  In order to measure the emotional reactions of participants towards the two different environments, the semantic differential Pleasure, Dominance and Arousal scale was used. Mehrabian and Russell (1974) designed this widely used instrument to investigate how consumer behaviors are influenced by atmosphere (Foxall, 1997). This instrument is based on three dimensions to describe an individual's emotional responses to an environment: PLEASURE, AROUSAL and DOMINANCE (Mehrabian, 1996). The PDA model consists of three nearly independent dimensions: pleasure - displeasure, arousal - non-arousal, and dominance - submissiveness. Whereas the *pleasure-displeasure* dimension distinguishes between the positive - negative affective quality of emotional states, the *arousal-non-arousal* dimension refers to a combination of physical activity and mental alertness, and the *dominance - submissiveness* dimension is defined in terms of control versus lack of control. For each dimension a number of paired items is rated along a scale (see Appendix A: Pleasure - Dominance - Arousal Scale). These items were rated after participants had completed the experiment.
- *NASA Task Load Index (TLX))*
  The task load evoked by the performance of the assignments is measured by means of the NASA Task Load Index (Hart and Staveland, 1988). In order to measure mental workload, the TLX uses six bipolar scales to assess task load on six dimensions: MENTAL DEMAND, PHYSICAL DEMAND, TEMPORAL DEMAND, PERFORMANCE, EFFORT and FRUSTRATION (Rubio et al., 2004). The mean task load value represents how demanding the participants experience the execution of the tasks. The TLX was completed after the experiment.

[3] although the mood could also have changed after the experimental session this was not measured

◇ *(Dis)confirmation of expectations*
In order to assess to which extend user's expectations are confirmed during the experiment, a commonly applied seven-point semantic differential scale is used (Aiello et al., 1977; Linda and Oliver, 1979; Oliver, 1977; Swan and Trawick, 1980; Westbrook, 1980). The scale ranges from *The experimental system was worse than expected* to *better than expected* and is completed after the experiment. In order to measure participant‘s expectation level (Churchill and Surprenant, 1982) seven-point semantic differential scale was used. The questions used from the (dis)confirmation of expectations questionnaire are:

My expectation about the environment were

Too high								Too low

My expectation about the system were

Too high								Too low

◇ *User Satisfaction Questionnaire (USQ)*
This is a questionnaire developed by De Ruyter and Hollemans (1997) which is aimed at measuring the user‘s perceived satisfaction regarding the usability of consumer electronics. Test participants were asked to complete this questionnaire after completing the actual test. The questionnaire consists of 54 items (see Appendix C: User Satisfaction Questionnaire) and measures the user's satisfaction on six dimensions:
Overall impression. The way the product presents itself to the user and the rising of expectations by the self-presentation
Efficiency. The expenditure of resources by the user in order to complete tasks while interacting with the product
Learnability. The degree to which a product facilitates learning by initial self presentation, the guidance of the user during the interaction and the answer given to questions about the product
Transparency. The degree to which the organisation of the visible parts of the product together with the workings make it clear how to interact with the product
Conformity. The degree to which a product complies to the expectations of the user regarding the organisation of the visible parts and use of indicators
Enjoyability. The degree in which the workings of a product stimulates

the use of the product and the degree to which the product looks good to the user

◇ *Expectations about the Environment*
Before participants would engage in the actual test one question with regard to their expectations of the testing environment was administered. Using a seven point scale the participants could express the level of their expectations:

My Expectations about the Environment are high

DISAGREE ☐☐☐☐☐☐☐ AGREE

◇ *Expectations about the System*
Similar, before participants would engage in the actual test one question with regard to their expectations of the experimental system was administered. Using a seven point scale the participants could express the level of their expectations:

My Expectations about the System are high

DISAGREE ☐☐☐☐☐☐☐ AGREE

Results

The data analysis of the results obtained with the different questionnaires is now presented in this section.

Expectations about the environment

A Mann-Whitney test was conducted on the level of *expectations* for the environment (as measured on a Likert scale) per testing *environment* for the first session [4]. The test did not yield a significant effect of *environment* on *expectation* scores (z = 0.83, P = 0.405).

Testing for a significant difference in the level of *expectations* of the environment between the first and second *session* (but ignoring for the type of *environment*) no significant effect is found (z = 1.31, P = 0.190).

There was no significant difference between the level of *expectations* for the *EARC environment* between the first and second *session* (z = -0.11,

[4] Since participants could have been biased from the first session, the second session was not included in this analysis

P = 0.912). However, a significant difference was found between the level of *expectations* of the *laboratory* between the first and second *session* (z = 2.99, P = 0.002). For test participants that would participate in a testing *environment* that is different for the first and second *session* (groups 2 and 3 in Table 4.1) there was no significance difference between the level of *expectations* for the *environment* between the first and second *session* (z = -0.22, P = 0.825).

When comparing the level of *expectation* of the testing environment measured before the second *session* [5] between participants that have been exposed to the *EARC environment* versus participants that have been exposed to the *laboratory environment*, no significant effect is observed in the *expectation* scores (z = -1.00, P = 0.317).

Expectations about the system

A Mann-Whitney test was conducted on the level of *expectations* of the experimental system (as measured on a Likert scale) per testing *environment* for the first session [6]. The test did not yield a significant effect of *environment* on *expectation* scores (z = -0.24, P = 0.810).

Testing for a significant difference in the level of *expectations* of the experimental system between the first and second *session* (but ignoring for the type of *environment*) a significant effect is found (z = -1.90, P = 0.005). There was no significant difference between the level of *expectations* for the *EARC environment* between the first and second *session* (z = 0.11, P = 0.912). No significant difference was found between the level of *expectations* for the *laboratory* between the first and second *session* (z = -0.16, P = 0.161). For test participants that would participate in a testing *environment* that is different for the first and second *session* there was no significance difference between the level of *expectations* for the *system* between the first and second *session* (z = -1.86, P = 0.06).

When comparing the level of *expectation* of the experimental system measured before the second *session* [7] between participants that have been exposed to the *EARC environment* versus participants that have been exposed

[5] Since participant could have been biased from the testing environment used in the first session, only the second session is used in this analysis

[6] Since participants could have been biased from the first session, the second session was not included in this analysis

[7] Since participant could have been biased from the testing environment used in the first session, only the second session is used in this analysis

to the *laboratory environment*, no significant effect is observed in the *expectation* scores (z = -0.80, P = 0.423).

Brief Mood Introspection Scale

Since the BMIS is assessed by means of 16 items on a Likert scale, the median of the BMIS score per participant will be used as measure of central tendency in the data analysis. When comparing the test participant's *mood* score measured before being exposed to one of the two testing *environments* (ignoring for test session) no significant difference is found (z = -0.79, P = 0.429). Conducting the same test for the first session does not provide a significant effect (z = -1.65, P = 0.09) nor for the second session (z = 0.69, P = 0.490).

Pleasure, Dominance and Arousal scale

The median for the PLEASURE, DOMINANCE and AROUSAL scale is used as measures for central tendency in the data analysis. A Mann-Whitney test on the *Pleasure* scale per testing *environment* provides a highly significant effect (z = 4.14, $P < 0.0001$) with a median of 5.5 for the *EARC environment* and 4 for the *laboratory environment*. For the *Arousal* scale there was no significant effect of testing *environment* (z = 0.41, P = 0.681) nor for the *Dominance* scale (z = 0.17, P = 0.856).

NASA Task Load Index

Conducting a t-test on the effect of *environment* [8] on the *mental demand* scale, no significant effect is found (t = 0.3, P = 0.763). No significant effect was found for the influence of *environment* on *Physical demand* (t = 0.85, P = 0.406), on *Performance* (t = -0.82, P = 0.420), on *Effort* (t = -1.09, P = 0.289) and on *Frustration* (t = -1.35, P = 0.192).

(Dis)confirmation of expectations

A Mann-Whitney test on the effect of *environment* on the *expectations about the environment* yields a statistical significant effect (z = 3.26, P = 0.001) as does the effect of *environment* on the *expectations about the system* (z = -1.91, P = 0.005).

User Satisfaction Questionnaire

Using a t-test on the effect of *environment* on the USQ scales is tested for

[8] In this analysis only participants being exposed to both environments are taken into the analysis as paired data points

statistical significance with matched pairs (i.e. test participants that have been exposed to both testing environments). The results of this analysis are listed in Table 4.6.

USQ dimension	*t*	*P*
OVERALL IMPRESSION	0.56	0.580
EFFICIENCY	0.56	0.580
LEARNABILITY	-0.39	0.697
TRANSPARENCY	-0.45	0.657
ENJOYABILITY	-0.37	0.712

Table 4.6: significance tests per USQ dimension between testing environments

Discussion

While there is a growing interest in the use of Experience and Application Research Centers (EARC) for studying user - system interaction with applications of advanced technologies, there have been no studies reported in literature on the potential response bias induced by such testing environments. This is remarkable since the field of user - system interaction research has been building on methodologies borrowed from the psychological research domain (Carroll, 1997) while the notion of response bias and empirical research paradigms to avoid this effect, are a key element in the psychology research domain.

Under the impulse of Ambient Intelligence many Experience and Application Research Centers have been established. The main characteristic these environments share is their level of realism in simulating a real life environment. Additional, these environments offer a context for integrating technologies into one embedded environment while studying the usage effect of test participants. Although the advantages of these environments is clear, there might be a problematic situation with respect to the data collected in empirical studies. More specific, the data might be influence by respondent's response bias due to the testing environment. The experiment discussed in this chapter investigates the presence of a potential response bias induced by an EARC during evaluations with users of advanced technology applications. The main finding of this study is that there is no statistical significant difference between the usability measures obtained in the EARC and those obtained in a traditional usability laboratory. The study presented follows on, from an earlier unpublished experiment (Nieuwenhuizen, 2006), which similarly did not provide evidence of any response bias. It appears that conducting user evaluations in the context of an EARC does not bias subjective report as compared to evaluating the same system in a traditional laboratory environment.

The statistical data analysis of this experiment indicates no significant effects of the testing environment on several measures of usability as measured for interacting with an experimental system. While a significant higher level of experienced pleasure was found with participants during the usability test in the EARC environment, this effect did not result in significant difference for *user satisfaction* with the experimental system, significant different *workload* when operating the experimental system or significant different *expectations* of the experimental system in the different testing environments.

During the debriefing at the end of the second session, several participants explained that because of performing the video editing tasks and filling in questionnaires they had paid hardly any attention to the environment. These remarks suggest that when testing systems that are more integrated in the environment participants could be paying more attention to the environment. Consider for example the evaluation of voice controlled environments in which there is no single point of interaction but in which the user will interact with or through the environment as a whole. To extend the generalizability of the findings reported future research should replicate this experiment with such applications that are more tightly embedded in the physical environment of the testing laboratory. Nevertheless, the absence so far for any evidence of response bias that is caused by the environment constitutes an important argument towards the validity and reliability of empirical research in environments such as the EARC.

4.3 Conclusions

The impact of stimuli in an environment (and more specific their congruency) on the user experience has been documented in literature (Mattila and Wirtz, 2001). Within the marketing research literature it is acknowledged that consumer‘s experiences such as satisfaction are influenced by more than a single system property (Beckwith et al., 1978; Moore and James, 1978) and contextual characteristics such as in retail stores (Wu and Petroshius, 1987). For example, a consumer‘s positive attitude towards a specific brand can influence the consumer‘s appreciation of an aspect from a product produced by that brand [9]. In a controlled study Wirtz and Bateson (1995) found a halo effect [10] for end-user satisfaction induced by matching end-user expectations.

The study reported in this chapter tested the hypothesis the that experimental environment will not have a statistical significant influence on the data collected in this context when compared to a typical laboratory environment. Indeed, the study supports this hypothesis and enables us to conclude that an environment such as an Experience and Application Research Center does not statistically significant influence the end-user‘s satisfaction ratings and perceived workload as typically collected in usability studies. It should be noted though that the experimental setup in this study did not require any specific characteristics of an environment to enable specific end-user experiences. While this might be seen as a shortcoming of the study it should be clear that the reported study was not intended to highlight differences between the environments (i.e. an EARC environment being more suitable for user-experience elicitation than a typical usability laboratory) but rather to assess the potential of the differences between the testing environments to influence the collected data. From this perspective one can assume that differences between both environments would have been induced by end-user expectations rather than functional differences between the environments.

[9] this phenomenon can be understood as an effect of cognitive dissonance

[10] a halo effect is decribed as a *distortion of the consumer‘s perception of* (product/service) *attribute-specific properties* (Wirtz and Bateson, 1995)

Publications for chapter 5

De Ruyter, B., Huijnen, C., Markopoulos, P., and Ijsselsteijn, W. (2003). Creating social presence through peripheral awareness. In Stephanidis, C. and Jacko, J., editors, *Human Computer Interaction, Theory and Practice*, volume 3, pages 889–893. Lawrenece Erlbaum and Associates

DE RUYTER, B. 2005. METHOD OF AND SYSTEM FOR AUGMENTING PRESENTATION OF CONTENT , Patent EP1590958

5

Mediating Presence

While communication media such as e-mail, telephony, text messaging services for mobile phones, etc., are common, there is more to system-mediated communication than exchanging factual information. As discussed in the case study in chapter 3, people have a need to share appreciation and experience social affiliation. This chapter describes an empirical study into the potential of attaining social presence by maintaining peripheral (visual) awareness of a connected person or group of persons, outside the context of communication/information exchange tasks. It further assesses affective benefits that arise out of this interconnection and illustrate the positive impact of awareness on social interactions.

5.1 Awareness systems

The interest in awareness systems was spawned by early research in Media Spaces (Bly et al., 1993), where awareness was considered as one of the primary benefits of providing sustained audio-video links among co-workers. Since that influential work an ever increasing variety of systems has been discussed under the general description of awareness systems and the term awareness itself has been given numerous interpretations. Awareness systems are best described as a class of systems that help individuals or groups build and maintain a peripheral awareness of each other (Markopoulos et al., 2009). In a social context, interpersonal awareness can be considered as an understanding of the activities and status of one's social relations, derived from social interactions and communications with them. Technology mediated awareness is achieved not by direct interaction or sharing a physical space but by means of sharing contextual information with remote sites. AmI environments that adhere to the principles of embedding and system intelligence are well positioned to support awareness applications for social presence. However, such environments do not go beyond the capturing and rendering of contextual information and are as such not classified as actors but merely as mediators of social presence.

Awareness systems research concerns a range of contexts of use: at work (Dourish and Bly, 1992; Begole et al., 2002; Convertino et al., 2008), at home (De Greef and Ijsselsteijn, 2001; Mynatt et al., 2003; Metaxas et al., 2009) and on the move (Khan et al., 2009) and they may concern different time scales. Sometimes awareness can be beneficial for collaboration (e.g. in the work environment) and in others its main purpose is to support social interactions between connected parties (e.g. in the home environment). Awareness systems can take many forms and the various factors for the success of awareness system seem to reappear in different contexts. Some awareness systems rely on high bandwidth communication and media richness to connect people. In other cases it may be sufficient to convey a simple message "I am home" or "I am thinking about you" in subtle or even decorative form using purpose-designed information appliances. For an extensive overview of theories, applications and methodologies in awareness research see Markopoulos et al. (2009).

5.2 Social Presence and Group Attraction

Being aware of a remote context of a person you are connected with may elicit feelings of **social presence** and **group attraction**. The feeling of *social presence* can be defined as *the sense of being together and communicating with someone* (Ijsselsteijn et al., 2001). While, the main body of research on social presence has concerned its determinants and measurement of social presence, this study deals with the consequences and the nature of social presence in the context of the home.

In social life people are part of many groups that they interact with. The feeling of being a member of such group is called *group attraction*. In general a distinction can be made between two types of groups: primary and secondary groups (Cooley, 1909). Primary groups are small, close-knit groups such as families, friendship cliques, children‘s playgroups, emotionally close peers and neighborhoods. Secondary groups are larger and more formally organized and tend to be shorter in duration and less emotionally involving than primary groups. Examples of secondary groups are professional associations, business teams, and religious groups. From literature in psychology it is known that group membership influences thoughts, feelings and actions. Groups give a sense of belonging and worth to the individual (Tajfel, 1972). Group membership that connects us to others is the basis for our participation in social life.

In the present study we will focus on social presence and group attraction within primary groups. Primary groups are characterized by frequent face-to-face interactions, interdependency and strong group attraction among their members. According to Cooley (1909) primary groups are *fundamental in forming the social nature and ideas of the individual*. Earlier studies have shown that people desire more connections to family and close friends, e.g., see Hindus et al. (2001); Eggen et al. (2003); Khan et al. (2009). It is possible that members of such primary groups are distributed over different locations. With systems that are capable of enhancing social presence, we might be able to socially connect these different members of the group.

We anticipate that increased social presence will improve the relations between group members and that this will be reflected in the attraction individuals feel towards the group of persons they are interconnected with. Particular about this study is that we will use peripheral awareness (i.e. awareness information that puts a low demand on attention). The type of peripheral awareness information applied in this study does not require users to actively engage with the system. This approach was adopted as other studies had reported a negative effect of increased obligations as users had to actively engage with the awareness system (Kaye et al., 2005). While such an approach has been reported in literature in the Aroma (Abstract Representation of Presence Supporting Mutual Awareness) project (Pedersen and Sokoler, 1997) this study will investigate the potential for such systems to create feelings of social presence and group attraction. In the remainder of this chapter we report an empirical study that was conducted to validate this conjecture.

5.3 Silhouette visualization to create social presence

An experiment to study the effect of peripheral awareness on group attraction and social presence was conducted in HomeLab (see chapter 4 for a description of the infrastructure). More specific, the hypothesis that **silhouette representations of a person in a remote context will create a feeling of social presence with that remote person** is tested.

5.3.1 Participants

A total of 33 participants (all Dutch males), participated in the experiment. They were recruited as groups of friends who enjoy watching soccer games, who have no love relationships, as such relations could bias the measurements of group attraction (participants with a love relation would feel more attracted to each other than other group members). The friends were split (two persons in one location and one alone in another location) and placed in two different rooms. During the experiment all participants watched the same soccer game that was pre-recorded.

5.3.2 Experimental setup

A mixed experimental design was adopted. First, two kinds of viewers were distinguished: single viewer (only one person in a room) and group viewer (2 persons in a room). This was a between-subjects condition. Then the type of information offered about the remote site was implemented as a within-subjects condition, where each subject is exposed to every condition. The different conditions are the *control* condition, the *silhouette visualization* and a *full video* condition.

All participants watched the same pre-recorded soccer game [1]. In the control condition the persons could not see any visualization of the remote site but were told that their friends were watching the same match simultaneously. This condition provides a baseline to compare the different visualizations: it can very well be the case that people already experience a certain level of social presence when they know about each other that they are engaged in the same activity at a certain moment.

The second level of visual information was a black and white image displaying the contours of the people in the remote site that was up-dated in real time when there is a movement of the persons in that room. As such,

[1] although the soccer game was more than 30 years old (1974) it represented an important match in which The Netherlands lost the game from Germany

the visualization represents the silhouette of people in the remote site. This silhouette was projected on the wall above the TV screen presenting the soccer game in order to provide a peripheral view on the presence from the remote site.

In the third condition participants were shown a live video feed of the remote person(s) watching the soccer game. In this visualization, more detail is depicted and the people in the visualization are always visible. This is contrary to the silhouette visualization, where people only see silhouettes when there is a change in activity. The trials were counterbalanced to avoid any potential sequence effects. Every group watched the same match to prevent any effect from differences in games. The game was a classic game for the national Dutch team that can still be exciting for Dutch soccer enthusiasts, even though it is not a live broadcast. The argument for a pre-recorded game is that soccer games can vary in terms of their attractiveness while a pre-recorded game provides an equal stimulus for all groups.

In none of the experimental conditions did we use audio from the remote site as an experimental condition. From previous social presence literature it can be expected that adding audio to the visual channel would significantly increase the level of social presence (Short et al., 1976). In this study, however, it was not intended to achieve the maximum level of social presence or to assess the effect of different modalities on social presence. Rather, we were interested to explore minimal and undisruptive means for achieving social presence through peripheral awareness of a remote site using different visualization methods.

5.3.3 Measures

The independent variables for the experiment are the *amount of visual information* (none, silhouette or full video) and the *group setting* (single viewer or 2 persons together). Social presence was measured after each condition by use of the **IPO Social Presence Questionnaire** (IPO-SPQ) (De Greef and Ijsselsteijn, 2001) because the IPO-SPQ measures the user‘s sense of social presence. The IPO-SPQ makes use of two approaches to measure social presence. It uses the semantic differential items from Short et al. (1976) that measures more affective qualities of the medium (a Cronbach‘s alpha of 0.90 is reported for the semantic differentials). These items assess the user‘s comments on the media (in this study the visualization type). Next to these semantic differential items, the IPO-SPQ includes subjective attitude statements about the experience using a 7-point agree-disagree scale (a Cronbach's aplha of 0.72 is reported for the attitude statements). These items assess the social presence experienced by the users when engaged with the system. For the current experiment, two items were excluded from the IPO-SPQ that had previously been shown to score insufficiently in the reliability analysis (De Greef and Ijsselsteijn, 2001).

Another dependent variable is the level of Group Attraction as experienced by the subjects. This was measured by the **Group Attitude Scale** (GAS) (Evans et al., 1986) after each condition. Group Attraction is defined as: *an individual's desire to identify with and be an accepted member of the group*. The GAS is composed of an equal number of positive and negative statements.

In addition to the IPO-SPQ and the GAS measures, a set of additional questions was posed by means of a custom designed questionnaire. This set included items concerning involvement, interest, how distractive was the visualization, enjoyment, feelings of isolation, etc. Also, all experimental sessions were recorded on video for both locations (including the visualization at each location), allowing for detailed behavioural observation and coding of verbal and non-verbal behaviours. The results of the collected data are now discussed.

5.3.4 Results

Social Presence

On the basis of the data collected with the IPO-SPQ questionnaire, an increase in the amount of experienced social presence from the control, to the silhouette and the full video condition is observed (see figure 5.1). While the amount of experienced social presence between the single and group viewer is not significant different for the different visualization conditions, there is a significant effect of the visualization condition on the amount of experienced social presence ($F(2) = 119.2$, $P < 0.0001$). The social presence scores for the silhouette visualization are only marginally higher than the control condition in which no visualization of the remote communication partner(s) was present (see figure 5.1).

Figure 5.2 illustrates the results of the manipulations of the independent variables on the experienced social presence of the medium, as measured by the semantic differential scales of the IPO-SPQ, which are sensitive to various affective qualities of the medium. This effect of visualization type on the experienced social presence of the medium (see figure 5.2) is significant

($F(2) = 36.7$, $P < 0.0001$). The full video visualization was rated higher on social presence than the silhouette visualization (for both types of groups).

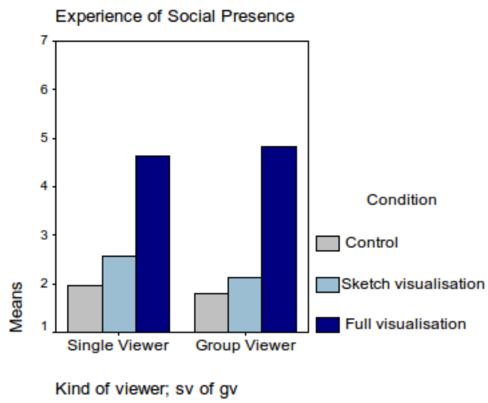


Figure 5.1: social presence per condition

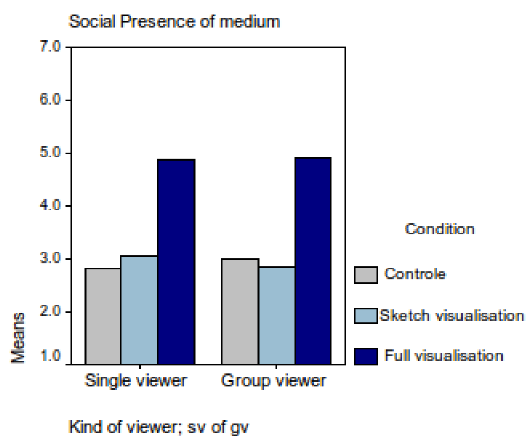


Figure 5.2: social presence of medium

Group Attraction

Figure 5.3 illustrates the results for group attraction as measured by the GAS. A significant difference can be observed between the results of the different types of viewers, with group viewers reporting higher group attraction than the single viewer in the control and silhouette conditions ($F(1) = 4.36$, $P < 0.045$). The effect of the full visualization on group attraction is the same for both kinds of viewers. Providing single viewers with full visual information about the rest of the group enhances their attraction to the group. The significant group attraction difference between the single and group viewer disappears when the silhouette or full video representation are displayed.

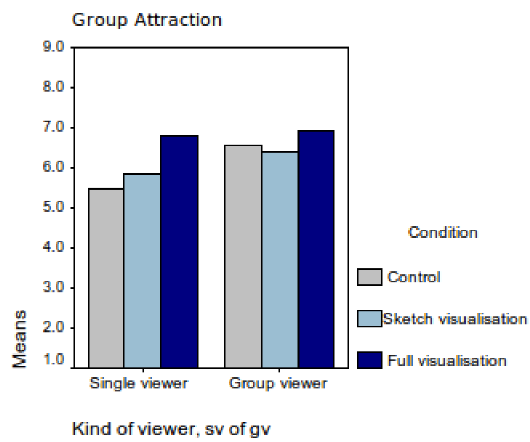


Figure 5.3: group attraction per condition

Interview data

Some items of the questionnaire gather information about the participant‘s attitude towards the experimental system. This information is obtained by asking the participants questions about the particular topic.

On the question whether the participants wanted to use the system, a difference was found between the conditions. The full video visualization scored remarkably higher than the silhouette visualization. Group viewers scored higher on wanting to use the full visualization at home than the single viewers. The interaction during the full video visualization was rated as much more fascinating for both type of viewers. As a potential negative effect, during the full video condition people felt they were being watched more than during both the silhouette and control conditions (see figure 5.4). There was no significant difference between group or single viewers.

To assess the disruptiveness of the different visual displays, we assessed the amount of user attention to the visualization during the experiment and the amount of user attention on the TV content during the experimental conditions. People devoted more attention to the television during the silhouette visualization than they did in the full video condition (see figure 5.5) while the

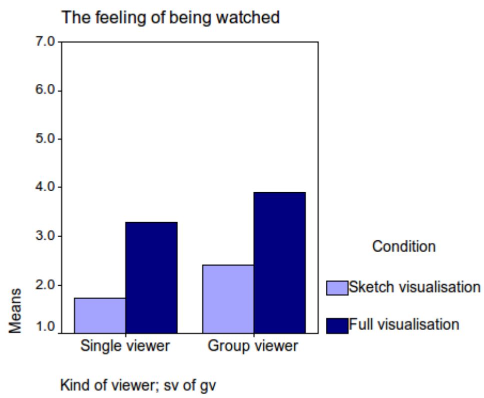


Figure 5.4: feeling of being watched per condition

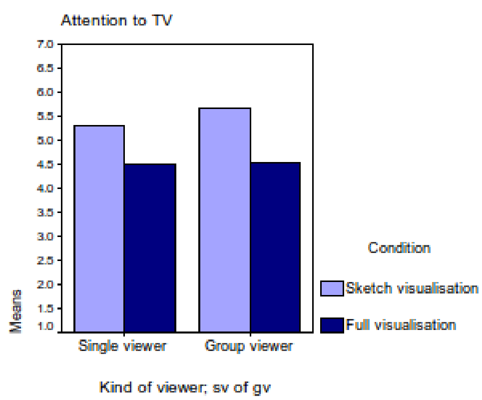


Figure 5.5: attention to TV per condition

silhouette visualization attracted less attention than the full visualization (see figure 5.6). Despite our assumptions regarding the calmness of the silhouette visualization, both the silhouette and the full video visualization were rated equally distracting; this was the same for both types of viewers (see figure 5.7). In line with this result, there was no difference in the attention group viewers reported to devote to their collocated friends (in the same room with them) across the different visualizations.

One of the motivations for choosing to study TV viewing as a shared experience for individuals separated by distance is that watching TV is to a large extent a social activity. This assumption was confirmed as most

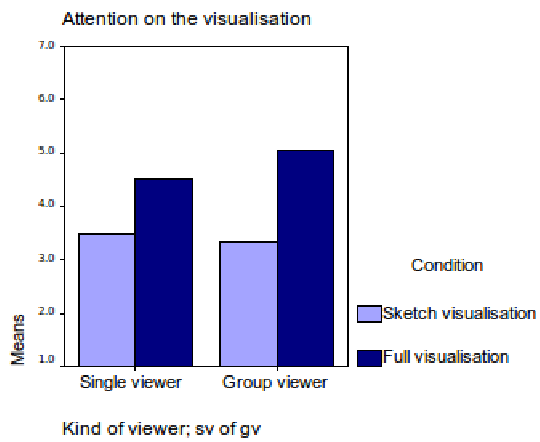


Figure 5.6: attention to visualization

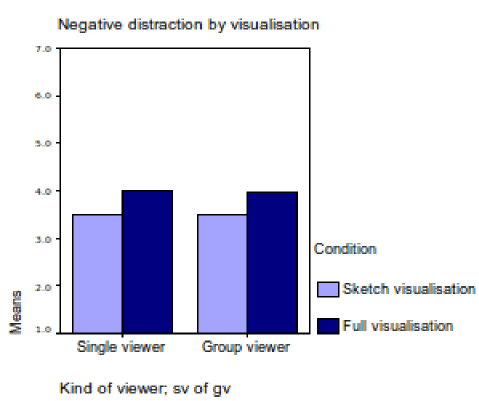


Figure 5.7: distraction by visualization

participants reported that they watch television together with at least one person, who is often their partner. Only 6% of the participants reported watching TV alone whereas 70% reported to watch TV with one additional person. Different motivations seem to underlie this tendency including the cosiness or atmosphere when you are together, and the possibility of discussing the contents of the programme together. Participants reported to prefer to watch sport and movies with others (they appreciate their reactions and company), whereas they prefer to watch news and documentaries alone. Additional, participants mentioned using the phone often while they are watching TV. It can be concluded that the participants want to communicate with others at remote locations more than they do now.

Most people indicated that a system that showed visual information about remote persons would enrich the interaction with their friends. Participants indicated they would like to have a system where they can communicate or interact with each other over distance while they are watching TV. They prefer to see and hear the persons. Most participants valued knowing the (emotional) reactions of the other person, and also wanted to know the opinion of the other person about the activity they are sharing. People‘s expressions were valued often as well, moreover, many participants liked to see what the environment looks like at the other person‘s end. Interestingly, group viewers thought that the single viewer had been feeling more left out and lonelier than was actually reported by the single viewer.

Observation data

The information that was obtained from observing the participants confirms the data of the questionnaires. The participants behaved in a way that seemed natural for a domestic environment, sitting relaxed in the couches with their drinks at hand. Group viewers verbally expressed that they considered the HomeLab living room as a real living room. Some of the group viewers remarked: *This is a very nice living room* or *I would like a couch like this at home*.

A recurring behaviour between collocated as well as remotely located parties that was seen in almost every session was making the sign cheers‘ with glasses. Another recurring behaviour was laughing and waving to the communication partner when a player of the opposite soccer team got tackled. During the silhouette visualization this was done as well, but not always recognized as such. This sometimes broke the action-reaction principle of communication, as during the silhouette condition such actions were not always identified as being an attempt to interact. This was different in the full visualization, where people were able to identify the behaviours and even the intentions of the interaction partner(s).

The urge to communicate with the remote partners was evident in many ways. In all three conditions, participants stated to their collocated friends that they missed audio. Some indicated that they would have liked to know sign language. Many group viewers talked about the fact that they wished they had access to their mobile phone that they had been asked not to use during the experiment. During the full video condition the exchange of messages was relatively easy since people were able to identify detailed movements, facial expressions and the context of the other location. In the full video condition two kinds of communications were prevalent. First, participants interacted with each other on the basis of events that happened in the game. Examples of this kind of communication are the waving to the other location when a football player missed a shot. Second, people initiated interaction with each other more or less independent of events in the game. Examples of this kind of communications were making signs of waving with hands to ask if the single viewer wants to join the group viewers, or indicating that the drinks are nice by smiling to the drinks.

Participants showed mirroring behaviour during the full video condition. For example, if the single viewer saw that the group viewers were laying back on the couch he sat relaxed as well; when whereas they moved to the edge of the couch, the single viewer bent forward often as well. Most of the time, in the silhouette condition only gross movements (like when a person is drinking) were recognized, and consequently most of the time they were based on a previous event in the game.

During the first couple of minutes participants tried out‘ the silhouette visualization to see what kinds of images were visible and how the other person(s) would react to their actions. It appears that during the silhouette condition the communications were mostly based on events in the game. In the control condition group viewers talked a lot about the single viewer. Group viewers often communicated that they felt sorry for the single viewer who was totally isolated. Group viewers often asked each other what the single viewer might be doing and mentioned that they missed the single viewer in the control condition. Overall, group viewers seemed more occupied with the single viewer than the single viewer was with the group viewers. It was also clearly visible that single viewers were more engaged in the game during the control condition. Single viewers were smiling less in the control condition.

5.4 Discussion

The present study investigated if social presence can be established by providing visual display of remote friends that are watching the same television program. By presenting different visualizations of the physical activities in the remote locations in parallel with shared TV content, an experimental condition was created in which the amount of social presence and group attraction could be measured.

In the full video condition people were constantly aware of the activities of their friends, whereas in the silhouette condition this awareness was not present all the time. The results from this study indicate that a low bandwidth visualization of the physical activities from remote locations is capable of establishing a sense of social presence. Furthermore, the feeling of being part of a group (i.e. group attraction) was increased. When the single viewers did not receive any representation about their friend‘s location, they had a low feeling of belonging to and being an accepted member of the group. The group viewer on the other hand, did have the feeling of belonging to the group in the control condition. When the minimal representation is introduced to the single viewer, this person did experience himself as being an accepted member of the group. The difference between the two kinds of viewers vanished by introducing the visual media. The silhouette visualization was perceived as equally distracting but the silhouette visualization gave participants not so much the feeling of being observed by the remote location and they reported paying more attention to the TV content. This latter aspect of the visualization could be of great importance to create social presence enabling systems for the home environment. Although the full video visualization was stronger in creating the feeling of being together and being part of a group, the low bandwidth visualization using silhouette representations of physical activity is probably more acceptable because it respects user privacy.

Further research needs to be conducted for establishing empirical evidence for the amount of social presence required: the present study did not consider answering the question if the amount of social presence created by the experimental visualization was sufficient to be of value for users. Test participants indicated that they would prefer different levels of social presence for different kind of TV content. People prefer to watch sports and movies in presence of others, whereas they prefer to watch news and documentaries alone. They do not want to be disturbed while concentrating on the serious programs. For entertaining programs, viewers enjoy making a cozy atmosphere and experience other person‘s reactions. More research is needed to investigate the context in which people prefer less or more social presence. Additional work is needed to estimate the generalization of research findings towards different relational types. It is very well the case that people prefer different levels of social presence depending on the interaction / communication partner.

The participants in this study were recruited on the basis of their strong personal relation. As such, a minimum feeling of being together and being part of a group was present. Moreover, different kinds of activities need to be investigated. It can very well be possible that different kind of activities elicit different levels of social presence.

5.5 Conclusions

The role and impact of awareness systems has been extensively studied in the context of working environments (Dourish and Bly, 1992; Begole et al., 2002; Convertino et al., 2008). In contrast to the workplace environment the focus for the home environment is on aspects such as creating feelings of intimacy (Mueller et al., 2005; Vetere et al., 2005) rather than optimizing collaboration between remote workers (Dourish and Bly, 1992; Begole et al., 2002; Convertino et al., 2008).

The present study focused on creating feeling of social presence and group attraction by sharing contextual information between two remote sites using different visualization techniques. By implementing the visualization techniques as peripheral awareness information we wanted to avoid the situation reported in Kaye et al. (2005) where users are obliged to communicate and cannot be engaged in a primary task such as watching TV. Our approach to designing an awareness system is in agreement with the approach Zhao and Stasko (2002) have adopted in designing a peripheral and unobtrusive awareness (and more specific the privacy preserving approach reported in Zhao and Stasko (1998)) display to feel connected to an on-line community. In contrast to the study reported by Zhao and Stasko (2002) however, there was an assessment of the perceived obtrusiveness (in terms of distraction and attention captured) of the awareness system in the study reported in this chapter. Interestingly, test participants wanted to engage with the remote site due to the awareness created by the different visualization techniques. Given the already social nature of watching TV (as reported by the test participants) and the clear desire to communicate while being engaged in this primary task, we observed that test participants were indeed eager to interaction (beyond awareness information) with the remote site.

The results of the study indicate that the social presence experienced by test participants increased with the richness of the awareness display (from silhouette to full visualization). This is in line with literature (Draft and Lengel, 1986) and demonstrates a challenge for designing awareness systems that also aspire for *calmness*. Important to note however is that in our study test participants reported feeling less watched by the silhouette visualization than the full visualization. This finding is particularly relevant as privacy concerns might be bottleneck for the widespread adoption of awareness systems.

Further research is needed to investigate what ingredients are needed to design *calm* awareness systems for creating feelings of social presence. It might be the case that showing the silhouettes that are visible continuously (as opposed to silhouettes that are based on movement) increases social presence; this has to be tested empirically. Also, abstract representations of activity, e.g. indicating a level of movement, might be an interesting option instead of the silhouettes shown in this study.

With respect to the manipulation of the group viewing (one person alone versus two persons in the same physical location) it is noted that there was no significant difference found between the social presence reported by the single versus the group viewers. This indicates that the single viewer feels as socially together with the group at the other location despite the physical distance between them and the isolation of the single viewer. When looking at the impact of the visualization technique it is clear that the full visualization is able to create a feeling of being together for the single and for the group viewers that is not significantly different between the viewers. Both kinds of viewers benefit equally from the richness of this medium.

Publications for chapter 6

De Ruyter, B., Markopoulos, P., Aarts, E., and Ijsselsteijn, W. (2004). Engineering the user experience. In Weber, W., Rabaey, J., and Aarts, E., editors, *Ambient Intelligence*, pages 49–62. Springer-Verlag

6

Mediating Availability

The experiment in chapter 5 reported some encouraging findings for using silhouette visualizations to create social presence between two remote sites. While it has to be acknowledged that a full video visualization is better for creating feelings of social presence, the silhouette visualization is less attention grabbing and is perceived as better shielding for privacy. Besides providing awareness information of the remote site, the silhouette visualization can also be used for interpreting the remote activities beyond getting a sense of awareness. Such interpretation relies however on empathic skills of the person in order to *live into* the remote situation since the awareness information (as implemented in our studies) is of low detail and only provides a vague impression of the remote context. Building on the important finding from the study reported in chapter 5 that such a level of awareness information is preferred in term of privacy above a richer full visualization approach, the studies in this chapter will build further on the silhouette visualization.

In Fogarty et al. (2004) a study is reported on the use of an awareness system for providing availability information. Interestingly the authors note that the awareness system is used for obtaining presence information rather than for making availability judgements. Based on these findings the authors conclude that providing such an awareness system will probably not reduce inappropriate interruptions as people do not use the presence information for making availability judgements. The studies reported in this chapter challenge this conclusion and further explore the role of an awareness system for mediating availability (for example to engage in a phone communication) of people in a remote context.

One of the essential aspects in making availability judgements based on awareness information is *empathy* as identification with and understanding of another‘s situation is required to make such availability judgements. During our social development it is essential that we develop empathic skills to exploit social relations in a manner that is acceptable for our social environment. Empathy is described as the intellectual or imaginative apprehension of another’s condition or state of mind without actually experiencing that person’s feelings (Hogan, 1969). However, empathy is also a mechanism that can be exploited in technology supported social interactions.

In this chapter we investigate how people use their empathic skills to interpret different awareness information visualizations (i.e. silhouette and full video visualization) representing situations in which a remote person is engaged. More specific the reported research investigates how people make judgments about the availability of a remote person based on several short video clips of typical home situations. Such video clips were presented to users of a home media space to help remote individuals or groups to maintain a sense of peripheral awareness of each other and, more specifically, to provide cues to each other regarding availability of others for communication.

6.1 Related work

Video has been used in a variety of awareness systems albeit most frequently in working environments where the video channel is used to create a media space for collaborative working (Boyle and Greenberg, 2005). These collaborative working environments are very different from home environments due to their well structured task setting. The use of video envisioned with the results of this study takes the form of *video-as-data* where the video channel is used to transmit contextual cues in a peripheral manner (Nardi et al., 1995) to create a sense of awareness. This sense of awareness can result in feelings of social connectedness or even trigger people to engage into social interactions such as phone conversations (De Ruyter et al., 2004). However, before engaging in a more active form of communication such as phone conversations, it is important to know if the remote person is available for this non peripheral communication.

Relying on people‘s empathic skills for making availability judgments is a fundamentally different approach than used in classical communication systems that require the receiver to indicate availability (or non availability) for communication. As discussed in Milewski and Smith (2000) there are two important challenges in designing awareness systems: (i) providing enough contextual information while respecting a person‘s privacy and (ii) not creating additional overhead for the end-user to control awareness information while providing the end-user a sense of control.

To overcome the overhead of setting availability information, the ConNexus system (Tang et al., 2001) automatically provides contextual cues to facilitate communication between people. However, the informativeness of this system is limited to device usage and location information of the remote person. In an attempt to find the balance between informativeness and privacy, Mayer et al. (2000) used a set of video snapshots of typical office situations for test participants to judge a remote person‘s availability to engage in an interaction. The results of this study provide insights into both the role of the presentation (i.e. picture resolution) and the content (i.e. the activity a remote person is engaged in) of video snapshots as basis for making availability judgments. The rationale for manipulating the picture resolution in this study was to find the minimal resolution suitable for availability judgments. The authors build on the assumption that lower picture resolutions of awareness information will provide greater privacy protection. Indeed, privacy is one of the major concerns in the use of video information (Boyle and Greenberg, 2005).

However, there are other techniques to shielding for privacy than reducing picture resolution: the study reported in chapter 5 has indicated that an abstract visualization technique such as silhouettes provides enough information to elicit a sense of social presence (i.e. the feeling of being together) while being preferred above full video visualizations in terms of shielding for privacy. In this chapter, we report on two experiments into the use of silhouette visualizations as awareness information for judging availability. While the first experiment seeks to discover potential interaction effects between empathy and availability judgments, the second experiments is aimed at further exploring these effects with a larger sample size and involves ground truth data for the availability.

6.2 Experiment 1

In this experiment the relationship between empathic skills and availability judgments of actors in video clips of typical home situations is studied. More specific the following two research hypothesis are tested:

1. **Silhouette visualizations are as good for making availability judgements as full video visualizations**
2. **Empathic skills of a test participant are related to making availability judgements**

In this study we are not per se interested in the accuracy of the availability judgements but on the impact empathy could have upon them and the role of the different visualization techniques.

6.2.1 Method

Participants

The experiment involved 44 voluntary and unpaid participants (with an average age of 35 years) that were recruited in both an academic and industrial research organization. Since the experiment was conducted using an online system, participants were invited via email and provided with an URL of the experimental setup. In total 26 males and 18 females completed the experiment.

Materials

The Empathy Scale (Mayer et al., 2000) was used to obtain a general empathy score from the participants. This 30 item questionnaire provides empathy scores that load on six sub-factors and a general empathy factor. While this instrument has been validated, it does not provide normative data. However, for our experiment this is not a problem since we want to compare empathy scores between experimental conditions (rather than making normative interpretations about the empathic scores). The contents of the video clips represent the following typical home situations: *Empty room*, *Leaving the room*, *Entering the room*, *Having dinner*, *Resting on the coach*, *Reading a magazine*, *Watching television*, *Making a phone call* and *Having a person to person conversation*. In contrast to the silhouette visualization used in chapter 5 the visualization in this experiment would always display the previous status of a context. This implies that even for the *Empty room* situation one would see the silhouettes of the context of an empty room.

While the *empty room* and *leaving the room* situations were chosen as a baseline for non-availability of the remote person to engage in a phone conversation with the participants, it is known that transition situations (e.g. entering or leaving the room) elicit clear availability judgments (Johnson and Greenberg, 1999). The other five situations are representative for a typical home context. The situations were presented as 20 second video clips with a 200x200 pixel resolution. Research has indicated that the minimum resolution for availability judgments based on video material should have a resolution of at least 128x128 pixels. The video clips could be either a full video visualization or silhouette type visualization. The silhouette visualization is realized by representing the context in the video by its contours only. The actors in the video clips are not related to any of the participants and there are no additional contextual cues such as time of the day available in the video clips themselves.

Experimental design

Data is collected using a 2x2 between-subjects experimental design. Participants complete the Empathy Scale in all conditions and then are assigned to one of the four conditions (11 participants per condition):

	Full video	**Silhouette video**
no contextual cues	video	silhouette
additional contextual cues	video + time of the day	silhouette + time of the day

Table 6.1: experimental design for experiment 1

In the first two conditions participants are asked to base their availability judgment either on full video visualizations or on silhouette visualizations of the same situations. In the second two conditions the participants would also get an indication of the time of the day in which this situation is taking place (marked as a text message above the visualization).

Procedure

After reading the test instructions and acknowledging the informed consent form, participants are asked to complete the Empathy Scale. Depending on the experimental condition, participants are asked to watch and rate on a five point scale (in terms of availability for engaging in a phone conversation with the persons in the video) a set of nine full-video or silhouette visualizations (see figure 6.1) with or without additional contextual cues (e.g. the time of the day). The order of the presented clips is randomized over participants. The clips start playing automatically and participants cannot proceed to the next clip without providing an answer to the availability question on a five point rating scale for the current clip. The availability question is phrased as: *After having seen the video, how would you judge the availability of the person(s) (based on their activity) to engage in a phone conversation with you?*.

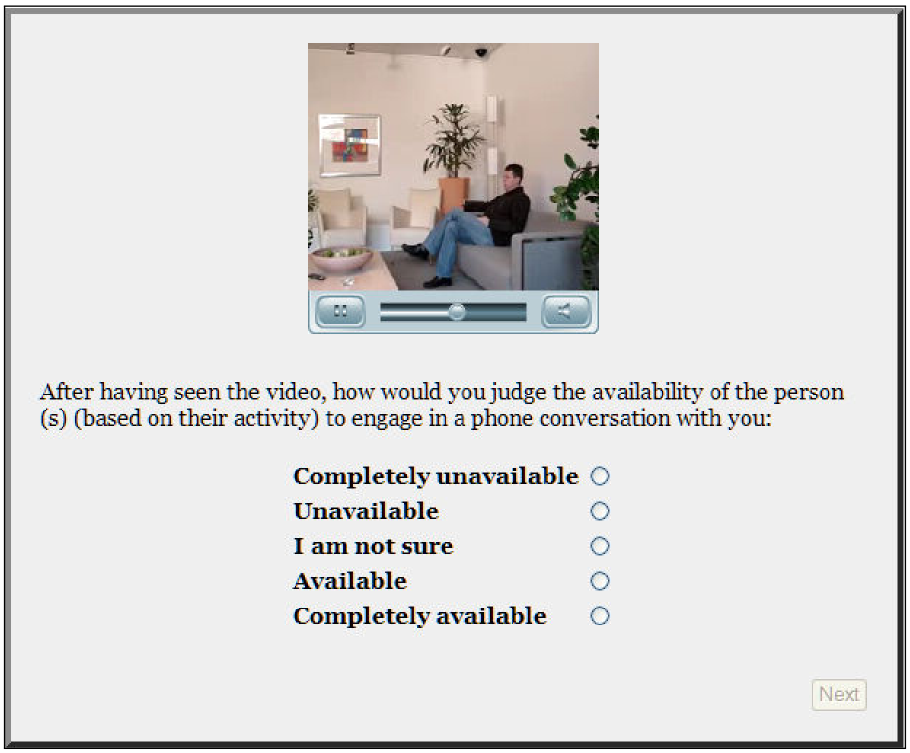


Figure 6.1: The video clips and availability rating

6.2.2 Results

Statistical data analysis is performed on the Empathy and Availability scores. When asked (after scoring the videos) to describe the situation represented in the full video or silhouette video[1], all participants (albeit in different wordings) correctly recognized the nine situations. Some participants also provided potential reasons for the context (e.g. *the person has gone to sleep* for the empty room situation).

[1] this was done to ensure that participants had a correct understanding of the contextual setting

Empathy scores

Using an unpaired *t*-test we found a significant ($P < 0.01$) difference in Empathy scores between male and female participants: female participants in this study demonstrated a significant higher empathy score than males.

Availability scores

A two-tailed Man-Whitney test (signed ranked sums) was used to test differences in availability judgments scores between the full video and the silhouette video visualizations shown to study participants.

- ◇ **Visualization type** Other than for the Eating ($P < 0.04$) and Leaving ($P < 0.01$) situation, there are no statistical significant differences in availability scores between the full video and silhouette video visualizations.
- ◇ **Contextual cues** When the additional context cue of time is given (as text above the video), there is a significant difference for the Away ($P < 0.01$) and Arriving ($P < 0.02$) situation. For all other situations there is no statistical significant difference in availability scores.

Availability by Empathy scores

Using the availability scores as predictors for the empathy score, we can perform a regression analysis on the obtained scores[2]. Since the experiment involved a small sample ($n = 44$) we used the adjusted *R*-square[3] for quantifying the amount of variance explained. Given the significant difference in empathy scores between male and female participants, we will conduct regression analysis for both groups separately.

- ◇ **Male participants** Considering the availability judgments for both the full video and silhouette video, we can explain 42% of the variance in the availability judgments from the empathy scores. When considering only the silhouette video a total of 82% of the variance in the empathy scores can be explained on the basis of the availability scores.
- ◇ **Female participants** Using availability scores for both visualization types together in the regression analysis, 55% of the availability‘s score variance is explained from the empathy scores. When using the silhouette video only data for females in this study into a regression analysis 98% of the availability‘s variance is explained.

---

[2] the use of regression analysis with ordinal level data such as obtained with Likert scales does not significantly affect Type I and Type II errors (Jaccard and W., 1996)

[3] the adjusted *R*-square value is calculated as $1 - (1 - R\text{-square}) * (n - 1) / (n - p - 1)$

6.2.3 Conclusion

The finding that contextual cues are important for the Away and Leaving situation seems logical since contextual cues such as time of the day give a completely different interpretation to the situation. Indeed, as some participants indicated: a person leaving the room late in the evening might suggest that this person is going to bed and meaning it is not such a good moment to give the person a phone call. Although there is no normative data available to interpret the Empathy scores, it is noted that female participants in this study demonstrated a significantly higher level of empathy than males. Also, for females there was found a stronger relation between empathy scores and availability judgments in the silhouette visualization condition. This effect was strongest for both males and females in the silhouette visualization condition.

The results of this experiment confirm the research hypothesis that the silhouette visualization is as good for making availability judgements as the full video visualization. Furthermore, the statistical analysis of the experiment data confirms the hypothesis that empathy plays an important role for making availability judgements.

6.3 Experiment 2

While the previous experiment provided insights into the relationship between empathy and availability judgments, it does not control for the correctness of the availability judgments. Given that the video clips in the previous experiment were staged by an actor, there was no criterion to judge the validity of the participants availability judgment. In this follow-up experiment we focus on the relationship between empathy scores and availability judgments that are controlled for their accuracy and seek confirmation for the potential effects found in the previous experiment.

In this experiment we are thus using *Empathic Accuracy* as independent variable. The term Empathic Accuracy (Ickes, 1993) was introduced to describe a person's ability to accurately infer what another person is thinking or feeling. In order to conduct a more reliable statistical analysis of the findings, this experiment will focus on the silhouette visualization condition only and involve significant more test participants in the silhouette visualization condition than in the previous experiment.

6.3.1 Method

Participants

For this experiment we recruited 46 voluntary and unpaid participants (with an average age of 26 years) in both an academic and industrial research organization. These participants did not participate in the previous experiment. Since the experiment was conducted using the same on-line setup (but with different content) as in the previous experiment, participants were also invited via email and provided with an URL of the experimental setup. Given that we did not want to test the gender differences for empathy we did not continue recruitment for achieving a more equal gender balance (as was done in the previous experiment).

Materials

While the same scenarios as used in the previous experiment are reused in this experiment, the videos are now recorded in real-life settings. Video recordings were made at different times of the day and week at the homes of 20 individuals as they go about their daily activities. From each individual a subset of clips were selected from the footage for which we obtained permission to use as stimuli in our experiment and for which we asked the person to rate their own availability for a phone communication at that moment when the video was recorded. For this, we used the same availability scale as in the previous experiment (this rating was done immediately after recording the video clip and served as ground-truth data for the experiment). As in the previous experiment these videos were transformed into silhouette representations to be used in the actual experiment. We selected video clips to represent the same set of situations as as in the previous experiment, only in this case they show actual situations and we have a ground truth regarding the availability of the person shown.

Experimental design

Since our main interest in this experiment is to assess the relationship between empathy and availability judgments in silhouette visualizations (controlled for the correctness of the judgment), the experimental design consists of only one condition. In this condition test participants will watch each silhouette video clip and make a judgment of the availability of the person in the clip (to engage in a phone conversation). In contrast to the previous study, the time of the day was not given in this experiment.

6.3.2 Procedure

The methodology adopted in this experiment follows the research methodology for measuring empathic accuracy as introduced by Ickes (1993). In this methodology a target person is being video recorded. At a specific point the recording is stopped and the target person is asked to provide what this person is feeling or thinking. In our study the question investigated was if the person would be available for engaging in a conversation with a remote person. During the study, the test participants watched the videos and were required to indicate if the target person would be available to engage in a conversation at the point where the video is stopped. The experimental procedure for data collection was similar to the procedure followed in the previous experiment and consisted of test participants going through the procedure for watching and evaluating video clips.

6.3.3 Results

The raw availability scores obtained from the experiment were corrected using the ground truth collected during the recording of the video clips. Although the availability scores are collected on a 5 point scale, we recoded the collected data to a correct or incorrect judgment by ignoring the spread over the extremes of the scale. In other words, we did not differentiate between a score completely available and available: both would be interpreted as available (both for the actors in the videos and the test participants) and similar for the unavailable scoring. So the data analysis was conducted on correct/incorrect scores based on the judgments available/don't know/unavailable. Further data analysis will be performed on the scores indicating the correctness of the availability judgment. The data was recoded accordingly because it would be rather strict to decide that an availability judgement would be wrong if test participants would for example provide completely available when the ground truth data was actually available.

Before conducting the statistical data analysis we verified the distribution of the Empathy scores in the data set (see figure 6.2) using the D'Agostino (D'Agostino et al., 1990) test for normality. This normality test is based on the skewness coefficient and indicates that our Empathy score distribution does not significantly deviate from a normal distribution ($P = 0.45$).

Similar to the data analysis in the previous experiment, we conducted a regression analysis. The regression analysis (based on $n = 46$) provided a solutions that explains almost 88% of the variance in the data set (adjusted

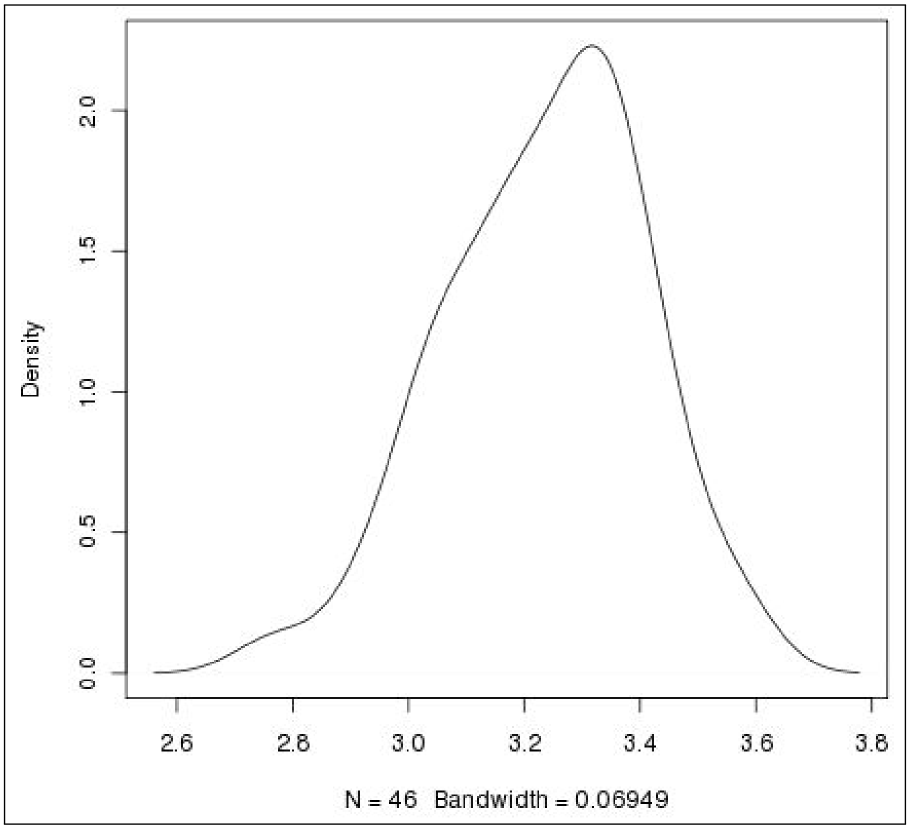


Figure 6.2: the spreading of the Empathy scores

R-square = 0.878). Furthermore, this regression solution is significant ($F = 37.91$, $P < 0.001$). The scenarios *Watching television* and *Having a person to person conversation* provided most correct availability judgments while *Having dinner* provided the least correct judgments.

6.3.4 Discussion

The results of experiment 2 partly confirm the findings of the first experiment: empathy plays an important role in making availability judgments from silhouette representations. Additionally, this experiment takes the correctness of the availability judgments into account: empathy scores are a good predictor of the correctness of the availability judgment made. This result is as we expected; observers with higher empathy are more capable of taking the perspective of the observed, and to evaluate whether this person is available for communication or social interruption or not. While expected, this result is quite novel in that previous research has not attempted to examine what individual characteristics influence availability judgment and influence the effectiveness of video obfuscation mechanisms. Further it points to the need for considering social skills of individuals and how to support them when designing applications for mediated social interaction.

6.4 Conclusions

The experiments provided evidence for an interesting trend in the relationship between empathic skills and availability judgments from awareness information. The analysis of the empathy scores of our participants provides findings that are in line with the earlier findings of Nicovich et al. (2005) indicating that women exhibit more empathy while interpreting awareness and presence information. Differences in empathic accuracy between women and men have been explained by various factors including awareness of the own empathic ability (Eisenberg and Lennon, 1983; Graham and Ickes, 1997), the gender of the target person (Klein and Hodges, 2001), the relationship with the target person (Simpson et al., 1995), the physical attractiveness of the target (Ickes et al., 1990).

As an alternative explanation of this effect Klein and Hodges (2001) have concluded that there is no difference in empathic ability between men and women but rather a motivational difference that accounts for different empathic ability exhibited. Once corrected for the motivational aspect, men exhibit as much empathic abilities as women. However, since our study did not introduce any mechanisms to influence the participant's motivation it is difficult to conclude on the underlying reason for the observed difference in empathy scores.

The general finding that empathic skills play an important role in making availability judgments from abstract awareness information such as silhouette visualizations is intriguing and can have an important impact on the way awareness and presence applications are designed. More specifically, our current study would suggest the use of contextual cues in awareness and presence applications that elicit strong empathic feelings with end-users. Although this requirement for awareness and presence applications has been followed implicitly by means of establishing a common ground between the connected users, it has not been articulated as a clear functional requirement.

From a design perspective the results presented indicate that higher priority should be placed on supporting through communication media people's empathic skills and, more generally, their social intelligence. A related design approach by Erickson et al. (2002) focuses on allowing people to carry over social skills from their daily life to their mediated social interactions. However, such research has not gone as far as to classify and describe concretely the social skills that one would expect to support in a particular situation. The current study would suggest the use of contextual cues in awareness and presence applications that elicit strong empathic feelings with end-users. Considering the design space of awareness systems (Markopoulos, 2009) this would favor using video imagery rather than abstracted information regarding individuals, it would favor expressiveness rather than just conveying factual information efficiently. Future investigations are needed for evaluating the relevant value of abstract information assembled through context sensing technologies versus the use of video media spaces in the home environment. We expect that they represent a different trade-off regarding privacy and awareness as discussed in the studies in chapter 5 and 6.

Publications for chapter 7

De Ruyter, B., Saini, P., Markopoulos, P., and Van Breemen, A. (2005). Assessing the effects of building social intelligence in a robotic interface for the home. *Interacting with computers*, pages 522–541

7

Mediating Social Intelligence

As discussed in chapter 2, the vision for Ambient Intelligence foresees that humans will be surrounded by embedded technology in the home. However, this will lead to an ever-increasing number of functionality that an end-user will need to know how to operate in daily life settings. This complexity gives rise to a necessary shift in the way users interact with this technology. There are two prevailing views on this communication (Markopoulos, 2004). One in which communication is through a multitude of task-specific information appliances (Norman, 1998). The other view is that a centralized user interface anticipates users' needs through an adaptive and intelligent system (Pentland, 2000). In the home domain, home-dialogue systems are expected to fulfill this latter role, acting as intermediaries between systems and services incorporated in the home and for the people in them.

There are several issues regarding the acceptance of home-dialogue systems by users. It is clear that users need to like and trust such systems in order for them to deliver their intended benefits. In this chapter, we explore the hypothesis that an appropriate design of the social behavior of these systems could lead to improved acceptance, not just of the home-dialogue system with which the user interacts but also of the embedded technology providing the functionality users are operating through that home-dialogue system. We will do so by designing a home-dialogue system that creates the perception of having social intelligence.

It is known that social intelligence enhances intercultural collaboration between humans (Dong and Collaco, 2009) but will social intelligence also enhance the human - system interaction when such system demonstrates social intelligence in the interaction? Since Reeves et al. (1996) published their work on the *media equation*[1] a lot of research has explored the notion that people react to media as if they were social actors.

There is a growing interest in on-screen agents and robotic characters that are endowed with human-like emotions and behavior. Some prototype systems have been made for the purpose of cultivating personal relationships (like friendship) with the user. They come with their own personalities, unique traits and functionalities. Others are found in a public domain, guiding users through their realm (Bickmore and Cassell, 2001; Heylen et al., 2004), providing information on products, offering services like online ticket sales. Studies such as reported in Breazeal and Scassellati (2000) present attempts to implement various human-like‘ characteristics into a robot (dialogue systems, facial expression recognition and presentation, gesture generation and recognition, humor perception and production, etc.).

The behaviors that these social/emotional characters (screen agents and robots) display range from facial animation (including eye-gaze and head movements) (Bickmore and Cassell, 2001; Heylen et al., 2004; Bickmore and Picard, 2004; Sidner et al., 2004; Thórisson, 1997) to limb and hand gestures (Beskow and McGlashan, 1997; Bickmore and Cassell, 2001; Bickmore and Picard, 2004; Thórisson, 1997); full-body posture adjustments (Beskow and McGlashan, 1997; Bickmore and Cassell, 2001; Breazeal and Scassellati, 2000) and locomotion (Bruce et al., 2002; Hüttenrauch et al., 2002). The standard research paradigm has been to compare characters with such expressive behaviors with characters lacking these. In the majority of cases the superiority of the expressive character over its neutral counterpart is demonstrated in terms of characteristics such as likeability and trust (Beskow and McGlashan, 1997; Bickmore and Cassell, 2001; Thórisson, 1997), naturalness of interaction (Beskow and McGlashan, 1997), satisfaction with interaction (Beskow and McGlashan, 1997; Heylen et al.,

[1] the research reported and discussed in this publication suggests that humans communicate with media as if they were human

2004), ease of use (Heylen et al., 2004), efficiency of task completion (Heylen et al., 2004), closeness to human characteristics (Thórisson, 1997), appropriateness of movements (Thórisson, 1997; Sidner et al., 2004; Heylen et al., 2004), joint attention (Imai et al., 2001), attention getting and cooperation elicitation from strangers (Bruce et al., 2002) and invested effort in interaction task (Bartneck, 2003).

Common to the studies mentioned here is that test-users reacted positively towards more expressive and human-like characteristics. Nevertheless, such research is limited to displaying relatively low-level behaviors and assessing fairly direct effects of these behaviors. The research presented in this chapter examines social intelligence as a higher-level construct and its effects on end-users interacting with a home-dialogue system providing the perception of having a level of social intelligence.

7.1 Social Intelligence

When we consider human-to-human social interactions, we see that there are several characteristics that make certain individuals stand out and more liked than others, or which convey an air of trustworthiness, competence and dependability (Ford et al., 1983; Keating, 1978; Sternberg and Smith, 1985). This list is large and includes attributes like *being nice and pleasant to interact with*, *being on time for appointments*, *thinking and speaking before doing* or *being sensitive to other people's needs and desires*. These attributes, and many others, are part of the concept of social intelligence; they can be seen as a continuum with the socially intelligent person falling in the high range. In its broadest definitions social intelligence is *...a person‘s ability to get along with people in general, social technique or ease in society, knowledge of social matters, susceptibility to stimuli from other members of a group, as well as insight into the temporary moods of underlying personality traits of strangers* (Vernon, 1933). So the socially intelligent person has a better than average ability to judge other people‘s feelings, thoughts, attitudes and opinions, intentions, or the psychological traits that may determine their behavior. This judgment creates expectations on the observer‘s part about the likely behavior of the observed person. This in turn leads to adjustments of one‘s own behavior accordingly and appropriately. However, that appropriateness can only be judged when the social context is taken into account. In this sense, social intelligence is not merely something that goes on between two people in isolation, but contextual factors also come into play.

Clearly, a person needs at least some degree of social intelligence to make his or her way in the world. This applies equally to socially complex situations (like business meetings) as to seemingly simple chores like grocery shopping. A considerable amount of research literature has been published in the field of psychology on social intelligence. Only few studies such as Johnson (2003) have modeled social intelligence in human - system dialogue. While it would be feasible, for example, to draw up a list of behavior characteristics that would indicate system intelligence, where the situations in which it is manifested are more confined, and thus relatively easily defined, this would be a daunting task when it comes to social intelligence. The situations imaginable are much more varied and are subject to constant change as interaction takes place. The manifestation of social intelligence is therefore almost entirely dependent on the context, making it difficult to provide an exhaustive list of characteristics of social intelligence. As a consequence there are many ways in which we can make an interactive system exhibit social intelligence. The following sections describe how we designed and implemented a set of behaviors to make a home-dialogue system be perceived as socially intelligent.

7.2 Experiment 1

The purpose of the study is to examine the effects of social intelligence in a robotic home-dialogue system such as the iCat. Although there has recently been significant interest in social intelligence in the domain of computational and robotic characters, the benefits that such intelligence might bring in terms of a better human - system interaction, have not been demonstrated or clearly identified. This study aims to address this apparent shortcoming. Rather than focusing on the direct effect of one or two behaviors, we took a broader approach: a number of aspects of social intelligence were implemented in a robotic character.

An experiment was conducted to examine the effects of social intelligence; more specifically we addressed the following research questions:

- ◇ Will the level of social intelligence implemented in the home-dialogue system be perceived? This question is important for the validity of the conslusions drawn from the study.
- ◇ What is the effect of bringing the concept of social intelligence into a home-dialogue system on the perception of the quality of the interactive systems (other than the home-dialogue system) in the environment?
- ◇ Will acceptance of home-dialogue systems increase if the concept of social intelligence is implemented in these systems?

7.2.1 Materials

For this study we used a robotic character, with which test participants could interact, to display social behaviors. Although there are obvious benefits of screen characters in terms of implementation facility, robots have properties of their own that make them compelling as candidates for a home-dialogue system. One of the most obvious differences is the physical embodiment of a robot in the world. Bartneck (2002) found that there was no difference in how his test participants enjoyed interaction with a robot and a similar onscreen character, but there was a social facilitation effect. This study found that participants would put more effort into a negotiation task with a robotic character than with a screen character. Another benefit of robots in terms of the social side of the interaction with their users is locomotion (walking or rolling around). Instead of the user approaching the interface for contact, the interface can approach the user. Taking this a step further, the user does not have to focus on the robot throughout the interaction. S/He can walk around and the robot can follow or turn its head as an indication that it is still listening‘ or attending‘ to the user. This turning of the head to track the user can also be used to look at other objects for communication purposes. Gestures can be performed that will make communication less ambiguous than communication with a screen character. Another option that flows from the locomotion benefit is the transport and manipulation of small objects for the convenience of the user. Other interaction techniques also come into play that is likely to make the interaction more natural than with screen characters. The robot can be touched. And the different touches can signify different commands, requests etc. In short, robots can open up a large and interesting design space to facilitate a more natural form of interaction with their users. It should be noted, however, that even though locomotion provides an important benefit over screen characters, we did not use this property in the current studies. The main reason for this is because our goal was to create a socially coherent character that would be present peripherally[2] even when the participant's attention was not focused on the robot. For our study, locomotion was not a requirement as the robotic interface was deployed specific to a fixed context.

The home-dialogue system used in our study takes the form of an *interactive Cat*, or just iCat. The iCat is a research platform for studying social robotic user-interfaces (Van Breemen, 2004; Leite et al., 2008;

[2] as will be described in the study design it was not aim of the study to have explicit interaction with the robotic interface but rather to use this system as a supportive device when operating a consumer electronics device in the home context

Ham and Midden, 2009; Looije et al., 2006). It has a cat-like appearance because of the acceptance of cats in domestic environments. The iCat is a 38 cm tall user-interface robot that lacks mobility facilities. The robot‘s head is equipped with 13 standard R/C servos that control different parts of the face, such as the eyebrows, eyes, eyelids, mouth and head position. With this set-up we can generate many different facial expressions that are needed to create an expressive character. Figure 7.1 illustrates facial expressions for the basic emotions happiness, surprise, fear, sadness, disgust, anger and neutral. A camera is installed in the iCat‘s head that is used for different computer vision capabilities, such as recognizing objects and faces. iCat‘s foot contains two microphones to record the sounds it hears, perform speech recognition and to determine the direction of the sound source. Also, a speaker is installed to play sounds (WAV and MIDI files) and to generate speech. Furthermore, iCat is connected to a home network to control in-home devices (e.g. lights, VCR, TV, radio) and to obtain information from the Internet. Finally, touch sensors and multi-color LEDs are installed in the feet and ears to sense when the user touches the robot and to communicate further information encoded in colored light. For instance, the operation mode of the iCat (e.g. sleeping, awake, busy, listening, etc.) is encoded in the color of the LEDs in the ears.

Traditionally robots are controlled using feedback control techniques, which result in rather machine-like movements of the robot‘s body parts (e.g. constant velocities). This makes the perception of a robot more like a machine than an intelligent and socially conversant agent. In order to create a socially communicative robot, the robot should be controlled in way which makes its behavior credible. The notion of credibility means that the actions of the robot should be apparent and understandable to a user. Credibility can be achieved by applying animation principles to the behavior generation module of the robot. Van Breemen (2004) has shown that traditional animation principles can indeed be applied to robots to improve their credibility. For the iCat, a dedicated robot animation engine was developed by Philips Research to control all of the iCat movements and to generate socially intelligent behaviors. A dedicated tool, called the Robot Animation Editor, was developed to graphically design believable robot animations, such as eye-blinking, facial expressions, head movements, etc. (Breemen, 2004).

7.2.2 Measuring Social Intelligence

In order to verify whether one system or the other conveys the impression of social intelligence to its users (the first research question in this study),

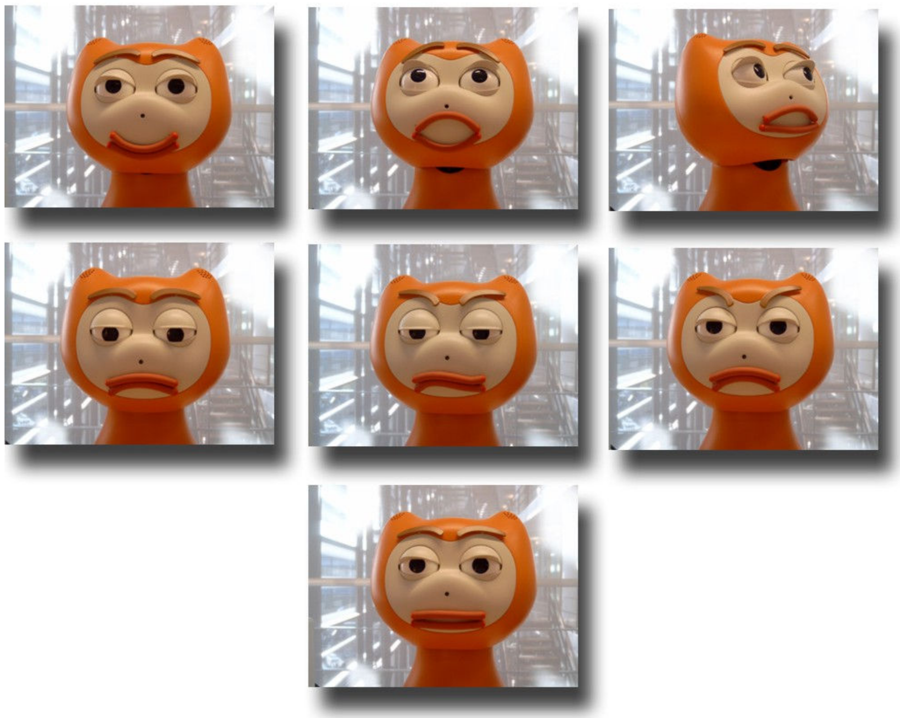


Figure 7.1: The iCat‘s facial expressions. From left to right: happiness, surprise, fear, sadness, disgust, anger, neutral

a measurement instrument was needed. In psychology, there have been different approaches to measuring the social intelligence of individuals. For example, photographs and stories have been used to see how well people perceive non-verbal cues and how they respond to moral dilemmas. But by far the most commonly used measurement instrument is the self reporting questionnaire. Although there are a number of good questionnaires assessing social intelligence (Gough, 1968; Strang et al., 1942; Banham, 1968), they seem to primarily be made up of questions one answers about oneself (self-reports). The current study calls for a measurement of the perception of the degree of social intelligence as perceived in an interactive system. Therefore we drew up a dedicated questionnaire, based on our knowledge of existing questionnaires. The first step in doing is so is establishing an item pool.

The online International Personality Item Pool (IPIP) questionnaire

(Goldberg, 1999) offers a wealth of questionnaire items that assess personality and social traits in oneself. It is a questionnaire that primarily reflects the big 5 dimensions that are also addressed by the Neuroticism-Extroversion-Openness Personality Inventory (NEO-PI) or the Neuroticism-Extroversion-Openness Five Factor Inventory (NEO-FFI) (Costa and McCrae, 1992). The difference is that the IPIP is in the public domain, and one is free to use any number of items to fit the purpose of the research being conducted. The IPIP items have the additional benefit that they have been compared and cross-validated with many existing questionnaires. A large number of items in the first version of the Social Behaviors Questionnaire (SBQ) were taken from the IPIP pool and then rephrased to address the perception of social intelligence in an animated character.

The procedure for choosing certain items rather than others was as follows. Items were selected on the basis of scales that in some way reflected affective responses to others, presenting oneself in a socially conventional way, and conveying an air of competence. Out of the item pool 172 items for 20 dimensions were selected (see figure 7.2). Items that were in the relevant scales were chosen on the basis of whether they made sense in the context of assessing an animated character, and whether the items could be answered by respondents with someone else in mind, rather than oneself. The result was an average of about 8-9 items per scale. Although a questionnaire with a total of 172 items is considered average in the field of psychology, it is considered fairly large in the area of human-computer interaction.

Altruism	Assertiveness	Competence
Compliance	Eagerness of effort	Good impression
Gregariousness	Likeability	Responsibility
Amicability	Compassion	Sociability
Dutifulness	Empathy	Socialization
Helpfulness	Modesty	Straightforwardness
Trust	Warmth	

Figure 7.2: the initial SBQ scales

All 172 items were formulated in the form of statements, like: *The character dislikes talking about himself*, or: *The character trusts what others say*‘ to be rated on a 5-point scale (1 = agree, 2 = somewhat agree, 3 = agree nor disagree, 4 = somewhat disagree, 5 = disagree) for expressing the extent of dis/agreement with the statement. To further refine the questionnaire, following statistical analysis are performed: (1) a Cronbach Alpha for assessing internal consistency of the items in the questionnaire and (2) a principal component analysis to assess if items would contribute to the expected dimensions of the questionnaire.

Data was collected using a one-factor (representing two levels of social intelligence) between subjects design. Participants were recruited and invited to go to a website where movies clips are presented. Participants saw either the movie clip in which the animated character acted in a socially intelligent way or the one where the character's conduct would be disastrous in terms of social intelligence. After viewing the video clip the participants were asked to fill in the questionnaire. They were instructed to answer the questions on the basis of the impression they had formed of the character in general. In practical terms this meant that the participants had to generalize, instead of just answer questions about the specific situation that was presented to them. In total, 82 participants took part in the online study for validating the SQB item pool.

The first analysis done was a Kruskal-Wallis one-way variance analysis to identify which scales the participants rated significantly differently in the two separate states. Out of 20 scales, 14 turned out to result in a significant difference between the two conditions (see figure 7.3). The overall mean of the scales also resulted in a significant difference. In short, there was a significant main effect of the condition on several scales, and the overall mean of the scales. The final version of the SBQ only considered the items from the scales that distinguished between the two states.

Next, the internal consistency of all 20 scales was measured by calculating Cronbach's alpha. In psychology, an internal consistency of 0.7 and higher is considered acceptable (Decoster and Claypool, 2004). On the basis of the alphas, it could be concluded that three-quarters of the scales were consistent enough to be used in the questionnaire. The scales *Compliance* and *Straightforwardness* resulted in a Cronbach's alpha that was considered too low for further consideration.

Finally, a principle components analysis was performed as an extra check to see whether the scales that were used had, in fact, some common underlying components. A five-factor solution emerged out of the analysis. When inspecting the items that loaded highly for the components it was possible to label them appropriately. Component 1 that explained 19.6% of the variance contained items that were indicative of someone *who cares* *for others*‘. The second component with 10.6% explanation variance can be termed *approaching others positively*‘. The third component *sure of own skills*‘ explains an additional 8%. The last two components *helpful*‘ and *cheerful*‘ explain 5.1% and 3% respectively, of the variance, resulting in a total variance of 46.3% explained. Intuitively, these components can be thought of as the main indications of what is considered socially intelligent. The components also mostly contain items in the scales that distinguished to the maximum the two states in the Kruskal-Wallis variance analysis.

Based on the findings from these statistical analysis the step was taken to reduce the number of items from 172 and 20 scales to make a reasonably‘ sized questionnaire that would distinguish between the two intended states to the maximum, as well as have scales that were consistent. The final version to be used in the study has 57 items (see appendix B) and 11 scales (see figure 7.3).

Altruism	Assertiveness	Competence
Sympathy	Sociability	Responsibility
Dutifulness	Empathy	Trust
Helpfulness	Modesty	

Figure 7.3: the final SBQ scales

Equipped with an instrument to measure the perceived social intelligence in a system (see Appendix B: Social Behaviors Questionnaire), an experiment to address the three research questions stated earlier, was conducted.

7.2.3 Method

This study was conducted in the HomeLab involving 36 participants who were recruited through an external agency and were paid for their participation. There were 15 women and 21 men. They all had at least some basic experience of using e-mail and the Internet.

7.2.4 Experimental setup

A one-factor between-subjects design was adopted with social intelligence being manipulated between conditions. Two conditions were created: A *socially intelligent*‘ state (condition 1) and a *socially neutral*‘ state (condition 2). Test participants were randomly assigned to one of these experimental conditions.

The experiment consisted of a number of tasks participants were required to perform (see tasks descriptions). Alongside, participants could interact with the iCat, that was standing on a table next to them and that acted as a supportive system for performing these tasks.

In both experimental conditions the robot spoke using synthesized speech from a text-to-speech engine. The speech was accompanied by lip-synchronization. During the *socially intelligent* condition the iCat blinked its eyes throughout the session, and displayed facial expressions and head movements. The behaviors in this condition were guided by a list of aspects indicative of social intelligence:

- ◇ Listening attentively: by looking at the participant when s/he talks and occasionally nodding its head.
- ◇ Being able to use non-verbal cues displayed by the user: responding verbally to repeated wrong actions on the part of the participant by offering help.
- ◇ Assessing well the relevance of information to a problem at hand: by stating what is going wrong before offering the correct procedure.
- ◇ Being nice and pleasant to interact with: by staying polite, mimicking facial expressions (smiling when participant smiles, for example), being helpful.
- ◇ Not ignoring affective signals from the user: by responding verbally or by displaying appropriate facial expression to obvious frustration, confusion, or contentment.
- ◇ Displaying interest in the immediate environment: the immediate environment being the participant and the equipment used in tasks, by carefully monitoring the person and the progress of the tasks.
- ◇ Knowing the rules of etiquette: by not interrupting the participant when s/he is talking.
- ◇ Remembering little personal details about people: addressing the participant by name, remembering login information, and passwords if asked.
- ◇ Admitting mistakes: by apologizing when something has gone wrong, but also when no help can be provided upon participant‘s request.
- ◇ Being expressive: by showing facial expressions while talking, if appropriate.
- ◇ Thinking before speaking and doing: by showing signs of thinking (with facial expression) before answering questions or fulfilling the participant‘s request.

The facial expressions and head movements were pre-programmed using the Robot Animation Editor mentioned earlier. The expressions were blocks of certain behaviors with different intensities. For example, there were four pre-programmed expressions for smiling‘, each at a different intensity. The experimenter would type in responses to the participant, although there were a few pre-set responses for situations that often occurred. For example, the instructions for working the DVD recorder (see tasks description) were pre-fabricated. The experimenter would initiate these pre-programmed social behaviors at appropriate moments during the sessions (Wizard of Oz). While the reported experiment uses a partial wizard-of-Oz technique (i.e. triggering the social intelligent expressions) others (Leite et al., 2009) have developed a framework for supporting a level of system controlled social intelligence for the iCat and techniques for automated affect recognition in test participants (Castellano et al., 2009).

Condition 2 was the *socially neutral*‘ state. In this condition the iCat did not display any facial expressions and did not blink its eyes. As in the *socially intelligent* condition, the iCat talked with used lip-synchronization. It responded verbally only to explicit questions from the participant. The aspects of social intelligence listed above, were not used in this condition. The only help provided without the explicit request by the participant would be in cases where they got stuck to such an extent that they would not be able to continue the experiment without help.

7.2.5 Experimental Tasks

Participants were asked to perform 2 tasks. The first task was to configure a DVD recorder for recording 3 broadcast shows for the upcoming week. This task was intended to let participants become familiar with having a robotic cat standing on the table that they could talk to in order to get support when operating the DVD recorder. The second task was an online auction system such as eBay but was special made for the purpose of this study in order to be able to manipulate the context. The task assignment was to buy several items at a given price from a list that was provided to them. At first, participants had to register as a new user using their actual email account. Participants were requested during recruitment to bring along their existing web-accessible e-mail credentials. During the experiment the participants would receive emails on items and bids in the auction system. A laptop computer was provided for accessing the online auction system and their personal email account.

Participants were informed that they could, at their own initiative, also give their e-mail credentials (login and password) to the iCat if they wanted it to take over the task of monitor the incoming messages related to the auction system. The rationale behind the choice of the auction task was that we wanted participants to be immersed in an intensive task, while in the background the iCat was always available for help. The task was also designed to convey the ability of the home-dialogue system to participate in the task by accessing sensitive information if authorized. Instead of artificial costs[3] we selected a task that required people to give the actual password of their e-mail system in order to let the iCat provide support during the auction task. This was done to make the task more personal and to assess the participant's commitment to take the risk of sharing confidential information to (and thus trusting) the iCat.

7.2.6 Measures

Several measures are used for data collection:

- ◇ *Social Behaviors Questionnaire (SBQ)* The newly developed SBQ was used to test whether participants rated the 2 experimental states as different. That is, if participants would perceive the presence of social intelligent in the experimental condition intended to provide this preception.
- ◇ *User Satisfaction Questionnaire (USQ)* This questionnaire has been developed at Philips Research and measures the user satisfaction with consumer products (De Ruyter and Hollemans, 1997). The USQ was used to assess the level of satisfaction with the DVD recorder participants had to operate during the experiment. Due to the effects of social intelligence in the larger cognitive context (including attitudes towards the DVD recorder), it is plausible that the quality of the interaction with the iCat may influence how participants experience the interaction with the DVD recorder. The specific research hypothesis to be tested here was the assumption **that satisfaction with the DVD recorder would be higher in the socially intelligent state**.
- ◇ *The Unified Theory of Acceptance and the Use of Technology (UTAUT)* . The UTAUT (Venkatesh et al., 2003) is a questionnaire to measure technology acceptance in the workplace. It is used to evaluate the likelihood that new technology introduced in industry will be accepted by the employees who will be required to use it. In our study the questionnaire was used to measure the extent to which participants would

[3]several experiments deploy the approach of giving an artificial budget participants have to handle but the actual cost and thus motivation to perform well is limited

use iCat at home after the experiment. Given the difference of the application domain, a revised version of the UTAUT was used. Again, because social intelligence was expected to affect more than just direct emotions, it was hypothesized that **technology acceptance would be higher in the condition where the iCat acted as socially intelligent**.

- ◇ *Rating of Own Performance* In a post-experimental questionnaire participants were requested to indicate on a five-point scale what they thought about their own performance during the experiment.
- ◇ *Observations* Finally, we noted the number of times that participants asked the robot general questions and the number of times they asked questions about the experimental tasks. We also noted the number of times that participants looked at the robot during the entire session or engaged in a conversation with the iCat.

This multiple set of measures was designed to test both the direct effects of iCat‘s behaviors on the participants and the potential implicit spillover effects like user satisfaction with the DVD recorder. These effects are considered spillover effects since the consumer device participants would interact with was the same for both experimental conditions. So if significant changes in user satisfaction would be found for the DVD recorder these would be spillover effects of the social intelligence in the iCat.

7.2.7 Procedure

Participants were welcomed and it was explained that they were going to perform two tasks. They were also told that while they did those tasks there would be a robot cat on the table next to them and that it could be addressed if they needed or wanted help. Additional, they were informed that there would be times that the robot would initiate conversation when it thought it might be able to help. This is all the information the participants were given about the iCat's role in the experiment. They were not informed about any functionalities that iCat had. The manner of instruction was such that emphasis was placed on completing the tasks with the DVD recorder and the online auction system, whereas interaction with iCat was secondary.

Next, the participants were given a task-booklet with instructions on the two tasks and were brought into the living room of the HomeLab. During the experiment they were left alone in the living room. They were given 10 minutes to program the DVD recorder to record three shows of their choice. The second task was the online auction task as described earlier. There was a laptop in the living room that had a broadband connection to the Internet.

Participants could access the auction by double-clicking a shortcut on the desktop.

Before bidding on items they were required to register as a new user with a valid web-accessible e-mail account. After that they were required to bid on and acquire several items. The participants were also told that simply bidding on an item would not ensure acquiring the item. Others' on the web were also bidding and they could be out-bid by these others. Notifications of higher bids were sent to the e-mail account that participants used to register at the start of the experiment. The e-mail account had to be monitored if the participants wanted to complete the task successfully. When the opportunity arose, the iCat could monitor their e-mail account for 'outbids' if the participants authorized it to do so. The iCat was there to help in many other ways as well. If participants could not manage to register as a new user, the iCat could register on their behalf. This was done in both states; in the socially neutral condition this was done only when participants did not succeed in registering within 12 minutes. The iCat could also give information on the items that were offered on the auction. If authorized, it could place bids for the participants. The participants were allowed 20 minutes to buy the listed items. The amount of items and bids was used to manipulate the pressure of the task (and thus create a potential situation in which the iCat could support them).

The sessions were recorded for further observational analysis. After performing the tasks participants were taken to a separate room were they gave their first impressions of the experiment in an interview. Then, they filled in the 3 questionnaires in the sequence SBQ, USQ and UTAUT. Finally, the participants were interviewed on their performance in the auction task.

7.2.8 Results from questionnaires

A Kruskal-Wallis one-way variance analysis were performed to see whether the responses to the 3 questionnaires differed from each other in the 2 experimental conditions. An $a = 0.05$ was set for all analysis.

SBQ: For the SBQ the difference between the two states was significant ($c2 = 5.938$, $df = 1$, $P < 0.05$). Inspecting the means indicated that our hypothesis was confirmed: participants thought the socially intelligent iCat was indeed more socially intelligent than the neutral iCat. Note that such difference is essential for the further analysis of the experiment since all hypothesis build on the assumption that both experimental condition differ in terms of demonstrated social intelligent behavior by the iCat.

USQ. The difference in evaluating the DVD recorder was also significant (c2 = 4.294, *df* = 1, $P < 0.05$). Participants who interacted with the socially intelligent iCat, were more satisfied with the DVD recorder.

UTAUT. The responses to the UTAUT also resulted in a significant difference (c2 = 9.633, *df* = 1, $P < 0.05$). Participants who worked with the neutral iCat were less inclined to want to continue working with iCat at home.

7.2.9 Results from the interviews

From the semi-structured interview it is clear that there was no difference between how participants evaluated their performance in the auction: on a scale from 1 (not pleased with performance) to 5 (very pleased with performance), the average for both states was 3.8 (c2 = 0.170, *df* = 1, $P = 0.680$).

At the end of the session people also had the opportunity to talk freely about the experiment. This also allowed each participant to answer the following question: *If you had iCat at home, what would you like it to do for you there?* There were responses like *all the electronics at home*'. Examples mentioned included the more obvious controlling the lights, heating system and entertainment equipment like TV, VCR and DVD player, but also less obvious things like the microwave, oven, and toaster were named. Some, however, also mentioned things like using iCat as a filter for Internet and TV programs for their children and having iCat as a cooking guide in the kitchen. More private tasks were also noted, like having their e-mail checked for them, screening telephone calls, and Internet banking. No differences were found in the pattern of responses of the participants in the different states. There was, however, a difference in the constraints imposed before authorizing iCat to access this personal information.

Of the 11 participants in the socially intelligent condition that would like iCat to handle their private tasks, 6 would like some additional research findings informing them that it is safe and secure to use iCat for these tasks. The other 5 would use iCat without further evidence. Participants in the neutral condition felt differently. There were 12 people who would authorize iCat for personal tasks. Four of them would like research evidence before using iCat. Only 1 person said he would let iCat deal with personal applications immediately. But 7 people were not sure if research would be enough. They wanted to experience iCat further before allowing it extended access to their private information. They stated that they would first give it small tasks. Only over the longer term, after proven success, would they give it full authorization. Only one of the participants in this condition would use iCat as it is now.

7.2.10 Results from the observations

We counted the times that participants asked iCat questions. The Wilcoxon-Mann-Whitney test was used to analyze the data. The averages in both states were close: 13.6 times in the socially intelligent versus 11.1 in the neutral condition (z = -0.954, $P > 0.05$). Additionally, up to 4.9 questions on average were posed about items in the auction in the socially intelligent state. In the neutral condition this average was 3.2 (z = -0.486, $P > 0.05$). Although at first glance there seemed to be a difference between the numbers of times that participants looked at iCat (11.6 and 6.0 for socially intelligent and neutral condition, respectively), it was not significant (z = -1.134, $P > 0.05$). In many cases participants looked at the iCat in anticipation of an answer.

7.3 Discussion

The results from the SBQ verify the distinctness of the experimental states that we wanted to create: participants rated the socially intelligent iCat as more social than the neutral one, which seems to validate that the collection of behaviors implemented or simulated in the iCat do help it to exhibit social intelligence.

The USQ also had a differential effect between the two experimental conditions. Since the USQ was developed to test satisfaction with a consumer product after thorough interaction with that product and the DVD recorder task only consisted of exploring one function in a time frame of 10 minutes, the significant difference found between the two experimental states is quite remarkable. Apparently the positive experience of the interaction caused spillover on the perception of working with the DVD recorder. It should be noted that for filling in the USQ, the participants were asked to ignore the help from iCat as much as possible: they were requested to strictly evaluate the DVD recorder.

Given the fact that we used a modified version of the UTAUT, we can only draw tentative conclusions from this measurement. Five out of the 6 scales were significantly different between the 2 conditions, indicating a positive attitude towards the social intelligent iCat.The UTAUT applied to the iCat and, as such, it shows the explicit positive effect of social intelligence manipulation.

There was no significant effect regarding perceived auction performance; most participants thought they did pretty well in both states. Being regular Internet users they were familiar with auction sites like eBay. The task therefore did not pose problems and they felt they did very well. In hindsight however, they indicated that they would have liked to ask the iCat more questions regarding the products in order to decide faster which products to bid on. The participants would also have liked to delegate more chores to iCat. Not all the participants, for example, discovered that they could ask iCat to keep an eye on their bids and have them notified when they were out-bid on an item. The 83% who did discover this function did delegate. Some participants tested iCat to the extent that they asked it to place counter offers when someone outbid them. They would also have liked iCat to give more reasons for recommending products.

The participants that were not very satisfied with how well they had performed were those who in their daily lives do not spend much time on the web or on the computer (5 participants). They had basic experience of using e-mail and the Internet, but were not as proficient as most of the other participants. Most likely, dissatisfaction was caused by the difficulty of navigating in unfamiliar territory, namely, an auction site.

The overall impression was that participants were themselves more *social*‘ with the socially intelligent iCat: they were much more inclined to laugh and divert conversation to a relational level. One participant, for example, was making conversation about the plasma TV set in the living room, commenting on how great it would be to watch soccer on such a cool screen, and whether iCat thought this living room would be available for rental that evening. Another person asked if iCat minded if he left the TV on to listen to some music while doing the task and continued by asking what type of music iCat enjoyed listing to. Participants also asked for more details on questions, than participants did with the neutral iCat. They were more curious about the reasons why the social iCat said the things it said than when they were interacting with the neutral iCat. For example, when asked which LCD monitor was a good one (to buy in the auction task) they were happy that iCat could help by naming a product. But they were curious how it knew this and why it was the best. They were also more inclined to ask iCat‘s opinion on the other LCD monitors. They asked these questions politely and using full sentences.

In the case of the neutral iCat, they were more inclined to take the suggestion of the best LCD monitor for what it was and not continue to probe further.

Participants in the socially intelligent condition liked the fact that the iCat was expressive in terms of facial expressions. They liked the fact that it nodded and shook its head in response to their talking. Overall, they agreed that it was a robotic cat with a face and it was only natural that it used its full potential this way. It made the iCat friendlier and easier to approach. On the other hand, participants in the neutral condition, who experienced only the talking and the lip-synchronization while talking, also liked iCat in that way. After all, they argued, it is a robot and it should not try or pretend to be anything other than that. Moving and facial expressions, according to them, would only look like a poor attempt to seem alive and it would likely annoy and distract you from whatever you are doing. This finding shows how hard it can be to imagine and evaluate something that has not been experienced.

Although many participants were reticent about disclosing their e-mail details to the iCat, many said that they would not have this problem at home. It would be a wonderful thought to not have to turn on the computer just to check e-mail, but that a robot could do that for you. At the very least, participants would like a robot to inform them whether or not they have new e-mail messages in their mailbox. For many of them, it is also important to know who the E-mails are from and the subject. Some even wanted to leave the reading of the e-mail to the iCat as well. Participants were also prepared to delegate other activities, like online banking to the iCat. Although they did expect to be kept informed about its activities. We note here that because of the experimental set-up, participants seemed to attribute quite extensive abilities to the iCat.

7.4 Conclusions

Studies on the effect of an embodied conversational agent (such as the iCat) capable of demonstrating emotional expressions have demonstrated encouraging results in terms of positive effects on the interaction in applications for special target groups like children (Markopoulos et al., 2008), elderly (Heerink et al., 2009) and children with autism (Dautenhahn and Werry, 2004). These effects include feeling more comfortable (Markopoulos et al., 2008; Heerink et al., 2009), richer verbal interaction (Heerink et al., 2009), more perceived ease-of-use (Heylen et al., 2004), improved learning in a cooperative human - robot learning situation (Hirohide, 2010) and even suggest a higher technology acceptance for these special target groups. Such higher acceptance was also found in the study reported in this chapter.

In contrast to these studies, the experiment reported in this chapter has adopted a well-defined approach for embedding social intelligence in a robotic interface. First, the type of social intelligent behaviors have been made explicit and implemented as part of the human - robot dialogue. Second, after developing a specialized questionnaire, the participant‘s perception of social intelligence in the robotic interface has been assessed in order to have a greater level of reliability of the experimental findings. Although our findings support earlier research findings from literature, they provide more reliable conclusions due to the more strict methodology adopted in our study.

Other studies have also demonstrated a spillover effect on the application for which the iCat is deployed. For example, Leite et al. (2008) have observed that demonstrating emotions and expressive behavior in an embodied conversational agent such as the iCat resulted in a better perception of a game that is being played with the iCat. Yet, the study reported in this chapter takes the next step in studying this spillover effect by observing this effect towards other technical systems in the environment. This spillover effect is considered as very powerful as it relates to a system (i.e. the DVD recorder) that has no functional nor technological relationship with the social intelligent system.

Based on the findings of our research we can conclude that embedding social intelligence in a robotic interface will:

- ◇ Enhance the end-user‘s acceptance of this robotic interface
- ◇ Create a positive spillover effect towards other systems in the environment

8

Conclusions

While one could be fascinated by techology and its developments, the applicability and relevance of this technology is not always obvious. Visionary researchers such as Weiser (1991) have been advocating scenarios of computational power being distributed on a large scale throughout an environment. It must be noted however that this vision on ubiquitous computing is technology driven and does not directly consider the application relevance of technology. Early 2000 the vision of Ambient Intelligence was introduced (Aarts et al., 2001) as a more human centric approach to the development of applications of technology. By focussing on the effect (or *user experience*) applications of technologies have on end-users, the AmI vision has triggered many futuristic scenarios to technology application and embedding into environments (ISTAG, 2001). While it is tempting to become fixated to some of these futuristic scenarios, one should use these scenarios not as a goal but as a means to application innovation. By doing so, the futuristic scenarios can be used as a tool for better understanding the potential consequences (both positive and negative) of innovation (Wright et al., 2008). Although AmI started of with a strong technological basis with applications mainly focussing on futuristic scenarios of home entertainment (e.g. the end-user experience of immersion in home entertainment systems), the vision has not been blind for societal trends such as the ageing population. Besides extending the vision towards futuristic scenarios of ambient assisted living, the extension of the intelligence model behind AmI has enabled the vision to deal with aspects such as social intelligence in AmI applications (as discussed in chapter 2).

The focus of this thesis is on social interactions mediated by technologies in an Ambient Intelligent environment. As discussed in chapter 1 as much as 20% of the population feels continuously socially isolated (Steffick, 1985) while it is acknowledged that deprivation from social connectedness can have a significant impact on both mental and physical well-being of human beings (Baumeister and Leary, 1995). From a install base perspective it is observed that network connectivity, the enabling condition for applications on social connectedness, is gaining ground with an estimated 63% of all households in 2013 having access to broadband internet (Nuthall, 2007). From an application perspective however, some major challenges such as *exceed basic user needs*, *appealing to people's interests* and *allowing people to customize* are still under investigation (Bodine, 2007). The studies reported in this thesis have provided valuable insights and a better understanding on how technologies can support social interactions by mediating presence, availability and social intelligence. The contribution of the reported work is on *infrastructures*, *methodology* and *applications* deployed in Ambient Intelligence research. These contributions are now discussed.

8.1 Statement of contribution

8.1.1 A Methodological Framework for AmI research

Positioning the end-user in the center during application development is a traditional approach in designing interactive systems (De Ruyter, 2003). However, designing interactive systems in the context of Ambient Intelligence raises some issues. First, the emphasis is not only on the usability of these systems but also on their potential to elicit specific experiences with end-users. Second, interactive systems can no longer be considered as stand-alone systems but are part of an environment populated with many different systems. The spillover effect investigated in chapter 7 provides a very powerful illustration of the influence of different systems in an AmI environment. Third, the vision of AmI requires embedding of technologies in such a way that these become *woven into the fabrid of daily life*. This embedding goes further than technical integration into an environment and requires a deep understanding of the contextual setting.

These requirements highlight the need for a methodological approach capable of integrating aspects such as *prevailing user needs*, *optimal contextual embedding*, *designing for usability and experience* and *impact on real life settings*. By adopting an approach that combines studies in *context*, *laboratory* and *field* different aspects of user experience research are taken into account. During contextual studies one can study aspects of the real life settings before new applications of technologies are introduced. Using a variety of ethnographic methods, rich contextual data is collected that can serve as a basis for not only understanding the future context of use but also the prevalling end-user needs that could be addressed with technological innovations. The laboratory environment offers a controlled context for assessing aspects such as initial end-user acceptance and usability of a proposed system. Through field studies one can gain insights into the effect and actual use of proposed systems in real life settings. Field studies are expensive in resources and will often be conducted with systems having limited functionality. However, all three phases of this research cycle provide their own specific findings (see table 3.1 in chapter 3) and will be addressed in an iterative manner.

8.1.2 Validation of EARCs as scientific instruments

One aspect of empirical research for AmI related application is the testing environment. In the 10 years of its existence we have observed the rise of Experience and Application Research Centers for investigating implementations of AmI scenarios. These EARC environments simulate real-life settings that are as realistic as possible. For example, in chapter 4 we discussed Home and Retail environments that are very realistic simulation environments of real-life settings. Such environments allow not only for the technical integration of applications of advanced technologies (that sometimes require more infrastructure than could have been installed in a real-life setting) but also support studying how end-users react and interact with these simulated environments. While within the research area of social sciences it is rather obvious to consider potential response bias by test participants due to the testing context, the potential influence of a testing environment on the acquired experimental data for EARC environments has not been assessed before.

In chapter 4 we report on a controlled usability study conducted in two types of testing environments: a traditional laboratory environment that does not simulate any aspect of a real-life setting and an EARC environment that simulates a complete home setting. This study revealed that traditional usability indicators such as perceived workload and user-satisfaction were not influenced by the testing environment. We did find a significant higher amount of pleasure test participants experience when interacting with an experimental system in an EARC context. Most important however is that such an higher amount of experienced pleasure does not influence the usability indicators that are of direct importance for evaluating an experimental system. As a first checkpoint this is an important finding since it allows the AmI research community to be confident in the results of their usability studies conducted on AmI type of systems.

8.1.3 Demonstration of the affective benefits of Awareness Systems

In chapter 5 and 6 several studies into the affective benefits of an awareness system have been assessed. In particular, we measured the level of social presence and group attraction elicited by an awareness system. While a richer awareness system such as full video visualization resulted in higher levels of social presence, it was shown that a less rich visualization technique such as silhouettes shields better for privacy and is still capable of supporting social presence and group attraction. Additional, the silhouette visualizations demonstrated being suitable for making availability judgements of the actors in a remote site.

The importance of empathy for making availability judgements based on awareness information was investigated in chapter 6. From these findings we can conclude that when designing awareness systems higher priority should be placed on supporting through communication media people‘s empathic skills and, more generally, their social intelligence. In other words, the research findings in this thesis emphasize the importance of designing awareness systems as social translucent systems (Erickson and Kellogg, 2000; Erickson et al., 2002).

8.1.4 Assessment of Social Intelligence in a Home Dialogue System

Many studies have investigated the effect of interactive systems equipped with the possibility to express emotions and some forms of social intelligence. However, up-to-date no instruments for the systematic assessment of perceived social intelligence in interactive systems have been established. In this thesis we report on a first step towards a standardized instrument for measuring perceived social intelligence in an interactive system.

8.1.5 Illustrated the spillover effects of Social Intelligence

In chapter 7 an interactive system, such as a robotic interface, was equipped with a level of social intelligent behavior. After verifying that this behavior was indeed perceived by the test participants, they interacted with another system in the same environment as the robotic interface. In the experimental condition where the robotic interface did exhibit social intelligent behavior participants were significantly more satisfied about the use of the other system. From this an important spillover effect from the social intelligence in one system towards other systems in the same environment was observed. Today‘s literature reports only few spillover effects and if so, these are related to at most the task for which the social intelligent system is used. This effect has not been reported for other, non-related systems, in the environment such as in the study reported in chapter 7.

These findings have far going consequences for the way interactive systems are designed for an Ambient Intelligent environment. These environments typically integrate many different systems and designing systems with the aim of delivering end-user experiences can thus not be done in isolation from other systems (and their potential influences) in the environment. The research findings presented in chapter 7 require us to think beyond the technological consequences of Ambient Intelligent environments (as has been mostly the case in the AmI field) towards the design of systems that will be positioned in such environments. While the first steps in formulating design principles for AmI environments (Obermair et al., 2006), design guidelines for designing experiential systems (Rozendaal et al., 2007) and methods such as the use of ambient narratives (Van Doorn et al., 2008) for designing Ambient Intelligent environments have been taken, more is needed to support the design of applications for supporting end-user experiences in AmI environments. As recognized by Streitz (2008) the ubiquitously available functionality in AmI environments will provide new forms of interaction, communication and collaboration. The reported studies provide significant scientific evidence to emphasize the need for more research into suitable design methods for AmI environments that take into account the experiential spillover effect that AmI systems might have in such environments.

8.2 Limitations and future work

The effect and experimental studies reported in this thesis have been designed in a scientifically falsifiable manner: based on clearly specified methodological designs and statistical data analysis a set of pre-defined hypothesis have been tested. From those a number of conclusions and implications have been formulated and discussed. Although the research reported in this thesis has provided some significant contributions to the field of social interactions in Ambient Intelligent environments, some limitations and future work can be formulated. These are now discussed.

8.2.1 Eliciting User Experiences in EARCs

The research findings from chapter 3 have provided a strong basis for accepting the scientific reliability of using Experience and Application Research Centers as context for empirical research. Given the high profile such EARC environments have it could have been expected that a level of response bias was raised with test participants. However, testing this research hypothesis, it was concluded in chapter 3 that no such effect is at play.

The limitation of the reported study however is that it was conducted in the context of a traditional usability test where test participants are engaged in an interaction with one specific system. This situation could be different for AmI systems that actually rely on an environment to provide the end user experience. In such situations it could be expected that the environmental settings of the EARC do have a stronger influence on the assessment of the end-user‘s experiences (beyond the usability of one focal system). The problem with this situation however is that it becomes difficult to differentiate between the system and the environment (which is exactly the direction AmI is proposing). Hence, in this situation it will become difficult to decide if the elicited end-user experience is a side effect of the environment or the direct consequence of interacting with the AmI environment.

Related to this issue is the question of the importance of an EARC for conducting AmI research. In other words: could one have elicited the end-user experience in a more traditional laboratory environment. While answering this question will pose a serious methodological challenge, one could argue that the elicitation of the end-user experience using a contextual setting is an inherent characteristic of an EARC and that environments capable of doing so qualify to be called and EARC.

8.2.2 Awareness Systems that become an actor and a mediator

The studies reported in chapter 5 and 6 have studied the effects of technology applications that act as a mediator for social interaction between people. These applications took the position of capturing and rendering contextual information over a distance with the purpose of strengthening the social connection between people. In chapter 7 a next step was taken and the effect of an interactive system becoming an actor was investigated. This study provided amazing findings on the effect of social intelligence in such an actor. Future research should investigate the potential and effects of combining both aspects: an AmI system that acts as both a mediator and an actor. For example: instead of conveying awareness information that could be used by end-users to make availability interpretations, an embodied conversational agent could make this decision on the basis of this awareness information. Doing so, an AmI system could take the next level in the model presented in figure 2.2 of chapter 2.

8.2.3 Social Intelligence in interactive systems

The study in chapter 7 reported on the effects of an embodied conversation agent demonstrating social intelligent behavior. Although these findings are strong and their impact for designing interactive systems are significant, one should be aware that there are still many challenges for bringing social intelligence to an interactive system (Green and De Ruyter, 2010). One aspect that is currently not experimentally manipulated is the kind of application for which an interactive system with social intelligence is used. It is acceptable to assume that for example healthcare related applications of such system are fundamentally different from an entertainment oriented application. Up-to-date however scientific evidence is lacking and until falsified we can generalise our scientific findings obtained in chapter 7. Nevertheless, additional controlled experiments are needed to gain better insight into this issue.

Social norms

As discussed in Erickson (2009) social intelligence involves adhering to social norms. These norms prescribe how behavior should be adapted to a context and can be seen as unwritten rules about the social *do's and don'ts*. An example from our iCat study is the situation in which the test participant is getting frustrated by not being able to handle the DVD recorder. In this situation the iCat should rather look sad than happy as the latter would be inappropriate in a cooperative context. In a competitive context such as gaming however it might be acceptable that the iCat expresses happiness when the test participant is experiencing difficulties in handling a system. Exhibiting the ‘right behavior‘ in the ‘right situation‘ is what makes the iCat‘s behavior social intelligent. The definition of ‘right behavior‘ and ‘right situation‘ is determined in social conventions. While this seems logic in human-to-human interaction it does imply that the design of interactive systems, such as the iCat, with levels of social intelligence requires adaptation to changing contexts and social norms (e.g. cultures).

Expectations

As reported in the work of Erickson and Kellogg (2000) on the design of social proxies, it can be expected that end-users will start making assumptions and interpretations about the behavior of the social proxies. As also discussed in Erickson (2009) these assumptions and interpretations are not necessarily always correct. This brings about yet another challenge for the design of social intelligent systems: while social intelligence in an interactive system can bring a level of predictability into the the human - system interaction (since the interactive system is now adhering to social norms) it can also lead to unrealistic expectations about this system. For example, as indeed observed in the social intelligence study reported in chapter 7 test participants started engaging on a very human-to-human like manner with the social intelligent iCat. While the benefits of designing social intelligence into interactive systems are clear, it is obvious that raised expectations might have harmful effects on the interaction if the system cannot live-up to these expectations. Hence, designing social intelligent systems will also require the system to be capable in dealing with raised expectations of end-users.

Personality versus Social Intelligence

While our research study focussed on the construct of social intelligence there are other psychological constructs that could have been implemented. One of these constructs is personality. Although there are multiple definitions of the construct of personality, a widely accepted definition is given by Child (1968): *personality refers to more of less stable, internal factors that make one person‘s behavior consistent from one time to another, and different from the behavior other people would have manifested in comparable situations*. This definition states that personality makes a person unique and different from others. This is fundamentally different from social intelligent intelligence were social intelligence is defined as ... *a person's ability to get along with people in general, social technique or ease in society, knowledge of social matters, susceptibility to stimuli from other members of a group, as well as insight into the temporary moods of underlying personality traits of strangers* (Vernon, 1933). One element that stands out in this definition of social intelligence is that this construct reflects adhering to social norms by which one can demonstrate suitable social behavior.

Although there are several studies that have investigated for example the effect of speech (Cozzolongo et al., 2007), animacy (Bartneck et al., 2007), affect (Saerbeck and Bartneck, 2010) and social behavior (Leite et al., 2008), no studies have compared in a controlled experimental manner the effect of personality versus social intelligence in an embodied conversational agent. Such study could indicate an important effect on the end user's acceptance of embodied conversational agents. In fact, one could formulate the research hypothesis that, in certain application contexts, end-users will not accept a technical system to take the position of being unique and different than others due to its personality while adhering to social norms by demonstrating social intelligent behavior might actually increase social acceptance. It would be scientifically valuable to assess effects such as the user's acceptance for the embodied conversation agent's behavior (such as reported by Ham and Midden (2009)) in experiments that explicitly manipulate personality and social intelligence.

Bibliography

Aarts, E. and De Ruyter, B. (2009). New research perspectives on ambient intelligence. *Journal of Ambient Intelligence and Smart Environments*, 1(1):5–14.

Aarts, E. and Eggen, J. (2002). *Ambient Intelligence Research in HomeLab*. Publishers, Eindhoven, The Netherlands.

Aarts, E., Harwig, R., and Schuurmans, M. (2001). Ambient intelligence. In Denning, P., editor, *The Invisible Future*, page 235250. McGraw Hill, New York.

Abowd, G. and Mynatt, E. (2000). Charting past, present and future research in ubiquitous computing. *ACM ToCHI*, 7(2):29–58.

Adler, A. (1938). Social interest: A challenge to mankind. In J., L. and Vaughan, R., editors, *Social Interest: A challenge to Mankind*. London: Faber and Faber Ltd.

Adler, M. and Fagley, N. (2005). Individual differences in finding value and meaning as a unique predictor of subjective well-being. *Journal of Personality*, 73:79–112.

Aiello, A., Czepiel, J., and Rosenberg, L. (1977). Scaling the heights of consumer satisfaction: An evaluation of alternate measures. In Day, R. L., editor, *Consumer Satisfaction, Dissatisfaction and Complaining Behavior*. Indiana University.

Ainsworth, M. (1989). Attachments beyond infancy. *American Psychologist*, 44:709–716.

Alperin, R. (2001). Barriers to intimacy an object relations perspective. *Psychoanalytic Psychology*, 18(1):137–156.

Antonucci, T. (2001). Social relations: An examination of social networks, social support, and sense of control. In Birren, J. E., editor, *Handbook of the psychology of aging*. San Diego, CA: Academic Press.

Avlund, K., Lund, R., Holstein, B., Due, P., Sakari-Rantala, R., and Heikkinen, R. (2004). The impact of structural and functional characteristics of social relations as determinants of functional decline. *Journal of Gerontology: Social Sciences*, 59B(1,):44–51.

Bakeman, R. (2000). Behavioral observation and coding. In Reis, H. and Judd, C., editors, *Handbook of research methods in social and personality psychology*, pages 138–159. Cambridge University Press.

Bakeman, R. and Casey, R. (1995). Analyzing family interaction: Taking time into account. *Journal of Family Psychology*, 9(2):131–143.

Bakeman, R. and Gottman, J. (1997). *Observing Interaction: An Introduction to Sequential Analysis*. Cambridge University Press, 2nd edition edition.

Baltes, P. and Baltes, M. (1968). Psychological perspectives on successful aging: The model of selective optimization with compensation. In Baltes, P. and Baltes, M., editors, *Successful aging: Perspectives from the behavioral sciences*. New York: Cambridge University Press.

Banham, K. (1968). *Social competence inventory for adults: a social competence inventory for older persons*. Durham: Family Life Publications.

Bartneck, C. (2002). *EMuu – An Embodied Emotional Character for the Ambient Intelligent Home*. Eindhoven University of Technology.

Bartneck, C. (2003). Interacting with an embodied emotional character. *Proceedings of the DPPI2003 Conference*.

Bartneck, C., Van der Hoek, M., Mubin, O., and Al Mahmud, A. (2007). Daisy, daisy, give me your answer do!: switching off a robot. *Proceedings of the ACM/IEEE international Conference on Human-Robot interaction*, pages 217–222.

Baumeister, R. and Leary, M. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117:497–529.

Beckwith, N., Kassarjian, H., and Lehmann, D. (1978). Halo effects in marketing research: review and prognosis. *Advances in Consumer Research*, 5:465–467.

Begole, J., Tang, J., Smith, A., and Yankelovich, N. (2002). Work rhythms: Analyzing visualizations of awareness histories of distributed groups. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW)*, pages 334–343. ACM Press.

Bell, R., Meiselman, H., Pierson, B., and Reeve, W. (1994). Effects of adding an italian theme to a restaurant on perceived ethnicity, acceptability, and selection of foods. *Appetite*, 22:11–24.

Bellotti, V. and Edwards, K. (2001). Intelligibility and accountability: Human considerations in context aware systems. *Human-Computer Interaction*, 16:193–212.

Bernhaupt, R., Obrist, M., Weiss, A., Beck, E., and Tscheligi, M. (2007). Trends in the living room and beyond. In Cesar, P., Chorianopoulos, K., and Jensen, J., editors, *Lecture Notes In Computer Science*, pages 146–155. Springer-Verlag.

Berscheid, E. (1985). Interpersonal attraction. In Lindzey, G. and Aronson, E., editors, *Interpersonal attraction*. New York: Random House.

Beskow, J. and McGlashan, S. (1997). Olga - a conversational agent with gestures. *Proceedings of the IJCA'97*.

Beyer, H. and Holzblatt, K. (1998). *Contextual Design*. Academic Press.

Bickmore, T. and Cassell, J. (2001). Relational agents: A model and implementation of building user trust. *Proceedings of CHI 2001*, pages 396–403.

Bickmore, T. and Picard, R. W. (2004). Towards caring machines. *CHI Extended Abstracts*, pages 1489–1492.

Bly, S., Harrison, S., and Irwin, S. (1993). Media spaces: bringing people together in a video, audio, and computing environment. *Communications of the ACM*, 36(1).

Bodine, K. (2007). *Desirable Online Experiences: Taking website beyond Useful and Usable*. Forrester report.

Boyle, M. and Greenberg, S. (2005). The language of privacy: Learning from video space analysis and design. *ACM Transactions on Computer – Human Interaction*, 12(2).

Breazeal, F. and Scassellati, B. (2000). Infant-like social interactions between a robot and a human caretaker. In Dautenhahn, K., editor, *Special issue of Adaptive Behavior on Simulation Models of Social Agents*.

Breemen, A. (2004). Animation engine for believable interactive user interface robots. *Proceedings of the International Conference on Intelligent Robotics and Systems*, pages 2873–2878.

Bruce, A., Nourbakhsh, I., and Simmons, R. (2002). The role of expressiveness and attention in human-robot interaction. *ICRA 2002*, pages 4138–4142.

Brumitt, B., Meyers, B., Krumm, J., Kern, A., and Shafer, S. (2000). *EasyLiving: Technologies for Intelligent Environments*, volume 1927/2000. Springer.

Caldwell, M. and Peplau, L. (1982). Sex differences in same-sex friendships. *Sex Roles*, 8:721–732.

Campbell, W., Krusemark, E., Dyckman, K., Brunell, A., McDowell, J., Twenge, J., and Clementz, B. (1982). A magnetoencephalogramphy investigation of neural correlates for social exclusion and self-control. *Social Neuroscience*, 1:124–134.

Carroll, J. (1997). Human – computer interaction: Psychology as a science of design. *Annual Review of Psychology*, pages 48–61.

Carstensen, L. (1995). Evidence for a life-span theory of socioemotional selectivity. *Current Directions in Psychological Science*, 4:151–156.

Castellano, G., Pereira, A., Leite, I., Paiva, A., and McOwan, P. (2009). Detecting user engagement with a robot companion using task and social interaction-based features. *Proceedings of the 2009 international Conference on Multimodal interfaces*, pages 119–126.

Child, I. (1968). Personality in culture. In Borgatta, E. and Lambert, W., editors, *Handbook of Personality Theory and Research*. Chicago, Rand McNally.

Churchill, G. and Surprenant, C. (1982). An investigation into the determinants of customer satisfaction. *Journal of Marketing Research*, 19:491–504.

Cohen, S. and Wills, T. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, 98:310–357.

Convertino, G., Wu, A., Zhang, X., Ganoe, C., Hoffman, B., and Carroll, J. (2008). Designing group annotations and process visualizations for role-based collaboration. In Liu, H., Salerno, J. J., and Young, M. J., editors, *Social Computing, Behavioral Modeling, and Prediction*, pages 197–206. Springer US. 10.1007/978-0-387-77672-9_22.

Cooley, C. (1909). *Social organization*. New York: Scribner.

Costa, P. and McCrae, R. (1992). *Revised NEO Personality Inventory and NEO Five-Factor Inventory*. Odessa, FL: Psychological Assessment Resources.

Cozzolongo, G., De Carolis, B., and Pizzutilo, S. (2007). Social robots as mediators between users and smart environments. *Proceedings of the 12th international Conference on intelligent User interfaces*, pages 353–356.

D'Agostino, R., Belanger, A., and D'Agostino, R. (1990). A suggestion for using powerful and informative tests of normality. *The American Statistician*, 44(4).

Dautenhahn, K. and Werry, I. ((2004)). Towards interactive robots in autism therapy: Background, motivation and challenges. *Pragmatics & Cognition*, 12(1):1–35.

De Greef, P. and Ijsselsteijn, W. (2001). Social presence in a home tele-application. *CyberPsychology and Behaviour*, 4(2):307–315.

De Ruyter, B. (2003). User centred design. In Aarts, E. and Marzano, S., editors, *The New Everyday: Vision on Ambient Intelligence*. 010 Publishers, Rotterdam, The Netherlands.

De Ruyter, B. and Aarts, E. (2010). Experience research: a methodology for developing human-centered interfaces. In Nakashima, H., Aghajan, H., and Augusto, J., editors, *Handbook of Ambient Intelligence and Smart Environments*. Springer.

De Ruyter, B. and Hollemans, G. (1997). Towards a user satisfaction questionnaire for consumer electronics: Theoretical basis. *Nat.Lab. Technical Note*.

De Ruyter, B., Huijnen, C., Markopoulos, P., and Ijsselsteijn, W. (2003). Creating social presence through peripheral awareness. In Stephanidis, C. and Jacko, J., editors, *Human Computer Interaction, Theory and Practice*, volume 3, pages 889–893. Lawrenece Erlbaum and Associates.

De Ruyter, B., Markopoulos, P., Aarts, E., and Ijsselsteijn, W. (2004). Engineering the user experience. In Weber, W., Rabaey, J., and Aarts, E., editors, *Ambient Intelligence*, pages 49–62. Springer-Verlag.

De Ruyter, B. and Pelgrim, E. (2007). Ambient assisted living research in carelab. *ACM Interactions*, 14(4):30–33.

De Ruyter, B., Saini, P., Markopoulos, P., and Van Breemen, A. (2005). Assessing the effects of building social intelligence in a robotic interface for the home. *Interacting with computers*, pages 522–541.

De Ruyter, B., Van Geel, R., and Markopoulos, P. (2009). Measuring the response bias induced by an experience and application research center. 5859:235–244.

De Ruyter, B., Van Loenen, E., and Teeven, V. (2007). User centered research in experiencelab. In Schiele, B., Dey, A., Gellersen, H., De Ruyter, B., Tscheligi, M., Wichert, R., Aarts, E., and Buchmann, A., editors, *Ambient Intelligence*, volume 4794 of *Lecture Notes in Computer Science*, pages 305–313. Springer.

De Ruyter, B. and Van Schijndel, C. (2004). Discovering behavioral patterns in the use of a universal remote control.

Decoster, J. and Claypool, H. (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.

DeLongis, A., Folkman, S., , and Lazarus, R. (1988). The impact of daily stress on health and mood: Psychological and social resources as mediators. *Journal of Personality and Social Psychology*, (54):486–495.

Dix, A., Finlay, J., Abowd, G., and Beale, R. (2004). *Human-computer Interaction*. Pearson Education.

Doherty, G., McKnight, J., and Luz, S. (2010). Fieldwork for requirements: Frameworks for mobile healthcare applications. *International Journal of Human - Computer Studies*, 68(10):760–776.

Dong, Q. and Collaco, C. (2009). Overcome ethnocentrism and increase intercultural collaboration by developing social intelligence. In *IWIC '09: Proceeding of the 2009 international workshop on Intercultural collaboration*, pages 215–218, New York, NY, USA. ACM.

Donovan, R. and Rossiter, J. (1982). Store atmosphere: An experimental psychology approach. *Journal of Retailing*, (58):34–57.

Dourish, P. (2007). Responsibilities and implications: further thoughts on ethnography and design. *Proceedings of the 2007 Conference on Designing For User Experiences*.

Dourish, P. and Bly, S. (1992). Portholes: Supporting awareness in a distributed work group. pages 541–547. ACM Press.

Draft, R. and Lengel, R. (1986). Organisational information requirements, media richness and structural design. *Management Science*, (32):554–571.

Eggen, J., Hollemans, G., and Van De Sluis, B. (2003). Exploring and enhancing the home experience. *Cognition, Technology and Work*, 5(1):44–54.

Eisenberg, N. and Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin*, (94):100 –131.

Erickson, T. (2009). social systems: designing digital systems that support social intelligence. *AI & Society*, 23:147–166.

Erickson, T., Halverson, C., Kellogg, W., Laff, M., and Wolf, T. (2002). Social translucence: designing social infrastructures that make collective activity visible. *Commun. ACM*, 45:40–44.

Erickson, T. and Kellogg, W. (2000). Social translucence: an approach to designing systems that support social processes. *ACM Transactions on Computer-Human Interaction*, 7:59–83.

Evans, N., , and Jarvis, P. (1986). The group attitude scale. *Small Group Behavior*, 17(2):203–216.

Fiori, K., Antonucci, T., and Cortina, K. (2006). Social network typologies and mental health among older adults. *Journal of Gerontology: Psychological Sciences*, 61B(1):25–3.

Fogarty, J., Lai, J., and Christensen, J. (2004). Presence versus availability: The design and evaluation of a context-aware communication client. *International Journal of Human-Computer Studies (IJHCS)*, 61:299–317.

Ford, M., , and Tisak, M. (1983). A further search for social intelligence. *Journal of Educational Psychology*, (75):196–206.

Foxall, G. (1997). The emotional texture of consumer environments: A systematic approach to atmospherics. *Journal of Economic Psychology*, (18):505–523.

Gaver, B., Dunne, T., and Pacenti, E. (1999). Design: Cultural probes. *Interactions*, 6(1):21–29.

Gerstel, N. and Gross, H. (1982). Commuter marriages: A review. *Marriage and Family Review*, (5):71–93.

Goldberg, L. (1999). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality Psychology in Europe*, 7:7–28.

Gough, H. (1968). *Chapin Social Insight Test Manual*. Palo Alto, CA: Consulting Psychologists Press.

Graham, T. and Ickes, W. (1997). When women's intuition isn't greater than men's. In Ickes, W., editor, *Empathic Accuracy*, pages 117–14. New-York: Guilford.

Green, W. and De Ruyter, B. (2010). The design and evaluation of interactive systems with perceived social characteristics: six challenges. *International Journal of Artificial Intelligence & Society*, 25(2):203–210.

Hakwkley, L. and Cacioppo, J. (2007). Aging and loneliness: Downhill quickly? *Currect Directions in Psychological Science*, (16):187–191.

Ham, J. and Midden, C. (2009). A robot that says bad!: using negative and positive social feedback from a robotic agent to save energy. *Proceedings of the 4th ACM/IEEE international Conference on Human Robot interaction*, pages 265–266.

Hart, S. and Staveland, L. (1988). Development of nasa task load index: Results of empirical and theoretical research. In Hancock, P. and Meshkati, N., editors, *Human Mental Workload*, pages 239–250. Amsterdam: North-Holland Press.

Heerink, M., Kröse, B., Wielinga, B., , and Evers, V. (2009). Measuring the influence of social abilities on acceptance of an interface robot and a screen agent by elderly users. *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology*, pages 430–439.

Helliwell, J., Huang, H., and Putnam, R. (2009). How's the job? are trust and social capital neglected workplace investments? In Bartkus, V. and Davis, J., editors, *Social capital: Reaching out, reaching in*. Northampton, MA, US: Edward Elgar Publishing.

Heylen, D., Van Es, I., Van Dijk, E., and Nijholt, A. (2004). Experimenting with the gaze of a conversational agent. In van Kuppevelt, J., Dybkjaer, L., and Bernsen, N., editors, *Natural, Intelligent and Effective Interaction in Multimodal Dialogue Systems*. Dordrecht: Kluwer Academic Publishers.

Hindus, D., Mainwarin, S., Leduc, N., Hagström, and Bayley, O. (2001). Casablanca: Designing social communication devices for the home. *Proceedings CHI 2001*.

Hirohide, U. (2010). Effect of social robot's behavior in collaborative learning. In *HRI '10: Proceeding of the 5th ACM/IEEE international conference on Human-robot interaction*, pages 195–196, New York, NY, USA. ACM.

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology*, (33):307–316.

House, J. and Kahn, C. (1985). Measures and concepts of social support. In Cohen, S. and Syme, S., editors, *Social support and health*, pages 83–108. Orlando: Academic.

Hüttenrauch, H., , and Eklundh, K. S. (2002). Observations from a long-term user study with a service robot. *Proceedings of the 11th IEEE International Workshop on Robot and Human Interactive Communication*, pages 158–163.

Ickes, W. (1993). Empathic accuracy. *Journal of Personality*, (61):587–610.

Ickes, W., Stinson, L., Bissonnette, V., and Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology*, (59):730–742.

Ijsselsteijn, W., Freeman, J., and De Ridder, H. (2001). Presence: Where are we? *CyberPsychology and Behavior*, 4(2):179–182.

Imai, M., Ono, T., , and Ishiguro, H. (2001). Physical relation and expression: joint attention for human-robot interaction. *Robot and Human Interactive Communication*, pages 512–517.

Intille, S., Larson, K., Beaudin, J., Nawyn, J., Tapia, E., , and Kaushik, P. (2005). A living laboratory for the design and evaluation of ubiquitous computing technologies. *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, pages 1941–1944.

Iqbal, R., Gatward, R., and James, A. (2005). Ethnomodel: An approach for developing and evaluating cscw systems. *Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, pages 633–638.

ISTAG (2001). *Scenarios for Ambient Intelligence*. European Commission.

ISTAG (2004). *Involving users in the development of ambient intelligence*. European Commission.

Jaccard, J. and W., C. K. (1996). *LISREL approaches to interaction effects in multiple regression*. Thousand Oaks, CA: Sage Publications.

Jackson, P. (2008). *Consumer Electronics Get Connected*. Forrester.

Johnson, B. and Greenberg, S. (1999). Judging peoples availability for interaction from video snapshots. *Proceedings of 2004 International Conference onProceedings of the Hawaii International Conference on System Sciences*.

Johnson, W. (2003). Interaction tactics for socially intelligent pedagogical agents. In *IUI '03: Proceedings of the 8th international conference on Intelligent user interfaces*, pages 251–253, New York, NY, USA. ACM.

Kaikkonen, A., Kekäläinen, A., Cankar, M., Kallio, T., and Kankainen, A. (2005). Usability testing of mobile applications: A comparison between laboratory and field testing. *Journal of Usability Studies*, 1(1):4–17.

Kang, Y., , and Ridgway, N. (1996). The importance of consumer market interactions as a form of social support for elderly consumers. *Journal of Public Policy and Marketing*, (15):108–118.

Kaye, J., Levitt, M., Nevins, J., Golden, J., and Schmidt, V. (2005). Communicating intimacy one bit at a time. In *CHI '05 extended abstracts on Human factors in computing systems*, CHI '05, pages 1529–1532, New York, NY, USA. ACM.

Keating, D. (1978). A search for social intelligence. *Journal of Educational Psychology*, (70):218–223.

Khan, V., Markopoulos, P., and Eggen, J. (2009). An experience sampling study into awareness needs of busy families. *Proceedings of the 2nd Conference on Human System interactions*, pages 335–340.

Khan, V., Markopoulos, P., Eggen, J., Ijsselsteijn, W., and De Ruyter, B. (2008). Reconexp: a way to reduce the data loss of the experiencing sampling method. *Proceedings of the 10th international Conference on Human Computer interaction with Mobile Devices and Services*.

Kidd, C., Orr, R., Abowd, G., Atkeson, C., Essa, I., MacIntyre, B., Mynatt, E., Starner, T., and Newstetter, W. (1999). The aware home: A living laboratory for ubiquitous computing research. pages 191–198.

Kiecolt Glaser, J., Ricker, D., George, J., Messick, G., Speicher, C., Garner, W., and Glaser, R. (1984). Urinary cortisol levels, cellular immunocompetency, and loneliness in psychiatric inpatients. *Psychosomatic Medicine*, (46):15–23.

Kientz, J., Patel, S., Jones, B., Price, E., Mynatt, E., and Abowd, G. (2008). *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, pages 3675–3680.

Klein, K. and Hodges, S. (2001). Gender differences, motivation and empathic accuracy: When it pays to understand. *Personality and Social Psychology Bulletin*, (27):720–730.

Kotler, P. (1974). Atmospherics as a marketing tool. *Journal of Retailing*, (49):48–64.

Kraut, R., Scherlis, W., Mukhopadhyay, T., Manning, J., and Kiesler, S. (1996). The homenet field trial of residential internet services. In *Communications of the ACM*, volume 39, pages 55–63. ACM.

Kuhn, T. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.

Law, E., Vermeeren, A., Hassenzahl, M., and Blythe, M. (2007). Proceedings of the workshop towards a ux manifesto. *Proceedings of HCI Conference 2007*.

Leite, I., Castellano, G., Pereira, A., Martinho, C., Paiva, A., and McOwan, P. (2009). Designing a game companion for long-term social interaction. *Proceedings of the international Workshop on Affective-Aware Virtual Agents and Social Robots*, pages 1–3.

Leite, I., Martinho, C., Pereira, A., , and Paiva, A. (2008). icat: an affective game buddy based on anticipatory mechanisms. *Proceedings of the 7th international Joint Conference on Autonomous Agents and Multiagent Systems*, 3:1229–1232.

Li Jiang, L., Liu, D., and Yang, B. (2004). Smart home research. *Proceedings of 2004 International Conference on Machine Learning and Cybernetics*, 2:26–29.

Linda, G. and Oliver, R. (1979). Multiple brand analysis of expectation and disconfirmation effects on satisfaction. *Proceedings of the 87th Annual Convention of the American Psychological Association*.

Looije, R., Cnossen, F., and Neerincx, M. (2006). Incorporating guidelines for health assistance into a socially intelligent robot. *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication*, pages 515–520.

Magnusson, M. (1996). Hidden real-time patterns in intra- and inter-individual behavior: description and detection. *European Journal of Psychological Assessment*, (12):112–123.

Markopoulos, P. (2004). Designing ubiquitous computer human interaction: the case of the connected family. In Isomaki, H., Pirhonen, A., Roast, C., and Saariluoma, P., editors, *Future Interaction Design*. Springer.

Markopoulos, P. (2009). A design framework for awareness systems. In Markopoulos, P., De Ruyter, B., and Mackay, W., editors, *Awareness Systems: Advances in Theory, Methodology and Design*, pages 49–72. Springer.

Markopoulos, P., De Ruyter, B., and Mackay, W. (2009). *Awareness Systems: Advances in Theory, Methodology and Design*. Human-Computer Interaction Series. Springer.

Markopoulos, P., Verschoor, Y., Ijsselsteijn, W., and De Ruyter, B. (2008). Testing interactive products with the robot intervention method. *Proceedings of the 7th international Conference on interaction Design and Children*, pages 117–120.

Maslow, A. (1968). *Toward a psychology of being*. New York: Van Nostrand.

Mattila, A. and Wirtz, J. (2001). Congruency of scent and music as a driver of in-store evaluations and behavior. *Journal of Retailing*, 77(2):273–289.

Mayer, J., Caruso, D., and Salovey, P. (2000). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, 24(4):267–298.

Mayer, J. and Gaschke, Y. (1988). The experience and meta-experience of mood. *Journal of Personality and Social Psychology*, 55(1):102–111.

McDonald, S., Monahan, K., , and Cockton, G. (2006). Modified contextual design as a field evaluation method. *Proceedings of the 4th Nordic Conference on Human-Computer interaction: Changing Roles*, 189:437–440.

Mehrabian, A. (1996). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology: Developmental, Learning, Personality, Social*, (14):261–292.

Mehrabian, A. and Russell, J. (1974). *An approach to environmental psychology*. MIT Press, Cambridge, MA.

Meiselman, H. (1996). The contextual basis for food acceptance, food choice and food intake: The food, the situation and the individual. In Meiselman, H. and MacFie, H., editors, *Food choice acceptance and consumption*, pages 139–263. Glasgow: Blackie Academic and Professional.

Metaxas, G., Metin, B., Schneider, J., Markopoulos, P., and De Ruyter, B. (2009). Awareness of daily life activities. In Markopoulos, P., De Ruyter, B., and Mackay, W., editors, *Awareness Systems*, Human-Computer Interaction Series, pages 351–365. Springer London. 10.1007/978-1-84882-477-5_15.

Meyer, F. and Vom Bögel, G. (2008). inhaus2 - intelligent construction site. *Fraunhofer-Institut für Mikroelektronische Schaltungen und Systeme*, Annual Report 2008:58–62.

Milewski, A. and Smith, T. (2000). Providing presence cues to telephone users. *Proceedings of CSCW*.

Moore, W. and James, W. (1978). A study of the amount of halo in the perceptions of automobiles. *Advances in Consumer Research*, 5:481–484.

Moss, F., , and Hunt, T. (1927). Are you socially intelligent? *Scientific American*, (137):108–110.

Mueller, F., Vetere, F., Gibbs, M., Kjeldskov, J., Pedell, S., and Howard, S. (2005). Hug over a distance. In *CHI '05 extended abstracts on Human factors in computing systems*, CHI '05, pages 1673–1676, New York, NY, USA. ACM.

Murray, S. and Holmes, J. (2006). Optimizing assurance: The risk regulation system in relationships. *Psychological Bulletin*, 132(5):641–666.

Mynatt, E., Rowan, J., Tran, Q., Abowd, G., Rogers, W., and Siio, I. (2003). Designing home appliances for older adults. *Cognitive Studies*, 10(3):343–352.

Nardi, B., Kuchinsky, A., Whittaker, S., Leichner, R., and Schwarz, H. (1995). Video-as-data: technical and social aspects of a collaborative multimedia application. *Computer Supported Cooperative Work*, 4(1).

Neustaedter, C., Brush, A., , and Greenberg, S. (2007). A digital family calendar in the home: lessons from field trials of linc. *Proceedings of Graphics interface 2007*, 234:199–20.

Nicovich, S., Boller, G., , and Cornwell, T. (2005). Experienced presence within computer-mediated communications: Initial explorations on the effects of gender with respect to empathy and immersion. *Journal of Computer-Mediated Communication*, 10(2).

Nieuwenhuizen, K. (2006). First investigation of the potential halo effect of high-end usability laboratories on user evaluation questionnaires.

Norman, D. (1998). *The Invisible Computer*. Cambridge, MA: MIT Press.

Nuthall, P. (2007). Western european residential broadband forecast: 2008 to 2013.

Obermair, C., Ploderer, B., Reitberger, W., and Tscheligi, M. (2006). Cues in the environment: a design principle for ambient intelligence. In *CHI '06: CHI '06 extended abstracts on Human factors in computing systems*, pages 1157–1162, New York, NY, USA. ACM.

Oliver, R. (1977). Effect of expectation and disconfirmation on postexposure product evaluations: An alternative interpretation. *Journal of Applied Psychology*, (62):480–486.

Pedersen, E. and Sokoler, T. (1997). Aroma: abstract representation of presence supporting mutual awareness. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '97, pages 51–58, New York, NY, USA. ACM.

Pentland, A. (2000). Perceptual intelligence. *Communications of the ACM*, 43(3):35–44.

Petit, C. and Sieffermann, J. (2007). Testing consumer preferences for iced-coffee: Does the drinking environment have any influence? *Food Quality and Preference*, pages 161–172.

Postma, C., Lauche, K., and Stappers, P. (2009). Trialogues: A framework for bridging the gap between people research and design. *Proceedings of Designing Pleasurable Products and Interfaces*, pages 25–34.

Prager, K. (1995). *The psychology of intimacy*. New York: Guilford Press.

Putnam, R. (1993). The prosperous community: Social capital and public life. *The American Prospect*, (13):35–42.

Putnam, R. (1995). Bowling alone: Americas declining social capital. *Journal of Democracy*, 6(1):65–87.

Putnam, R. (1996). The strange disappearance of civic america. *The American Prospect*, 7(24):7–24.

Reeves, B., , and Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge: Cambridge University Press.

Romero, N., Van Baren, J., Markopoulos, P., De Ruyter, B., and Ijsselsteijn, W. (2003). Addressing interpersonal communication needs through ubiquitous connectivity: Home and away. In Aarts, E., Collier, R., Van Loenen, E., and De Ruyter, B., editors, *Proceedings of the First European Symposium on Ambient Intelligence*, pages 419–429. Springer.

Rose, A., Shneiderman, B., and Plaisant, C. (1995). An applied ethnographic method for redesigning user interfaces. *Proceedings of the 1st Conference on Designing interactive Systems: Processes, Practices, Methods, and Techniques*, pages 115–122.

Rozendaal, M., Keyson, D., and De Ridder, H. (2007). Product behavior and appearance effects on experienced engagement during experiential and goal-directed tasks. In *DPPI '07: Proceedings of the 2007 conference on Designing pleasurable products and interfaces*, pages 181–193, New York, NY, USA. ACM.

Rubio, S., Díaz, E., Martín, J., and Puente, J. (2004). Evaluation of subjective mental workload: A comparison of swat, nasa-tlx, and workload profile methods. *Applied Psychology: An international review*, 53(1):61–86.

Russel, S. and Cousins, S. (2004). Ibm almaden's user sciences and experience research lab. *Proceedings of the ACM Conference on Human Factors of Computer Systems*, pages 1079–1080.

Saerbeck, M. and Bartneck, C. (2010). Perception of affect elicited by robot motion. *Proceeding of the 5th ACM/IEEE international Conference on Human-Robot interaction*, pages 53–60.

Schutz, W. (1958). *FIRO: A three dimensional theory of interpersonal behavior*. New York: Holt, Rinehart and Winston.

Schwartz, R. (1993). Managing closeness in psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*, 30(4):601–607.

Seeman, T. (2000). Health promoting effects of friends and family on health outcomes in older adults. *American Journal of Health Promotion*, (14):362–370.

Short, J., Williams, E., and Christie, B. (1976). *The Psychology of Telecommunication*. John Wiley and Sons Inc. and London.

Sidner, C., Kidd, C., Lee, C., and Lesh, N. (2004). Where to look: A study of human-robot engagement. *ACM International Conference on Intelligent User Interfaces*, pages 78–84.

Simpson, J., Ickes, W., and Blackstone, T. (1995). When the head protects the heart: Empathic accuracy in dating relationships. *Journal of Personality and Social Psychology*, (69):629–641.

Sleeswijk Visser, F., Stappers, P., Van der Lucht, R., and Sanders, E. (2005). Contextmapping: experiences from practise. *CoDesign*, 1:119–149.

Steffick, D. (1985). *Documentation of Affective Functioning Measures in the Health and Retirement Study*. University of Michigan.

Sternberg, R. and Smith, C. (1985). Social intelligence and decoding skills in nonverbal communication. *Social Cognition*, 2:168–92.

Strang, R., Brown, M., and Stratton, D. (1942). *Test of knowledge of social usage*. New York: Columbia University.

Streitz, N. (2008). From cognitive compatibility to the disappearing computer: experience design for smart environments. In *ECCE '08: Proceedings of the 15th European conference on Cognitive ergonomics*, pages 1–2, New York, NY, USA. ACM.

Swan, J. and Trawick, I. (1980). Consumer satisfaction with a retail store related to the fulfillment of expectations on an initial shopping trip. In Day, R. L., editor, *Consumer Satisfaction, Dissatisfaction and Complaining Behavior*. Indiana University.

Tajfel, H. (1972). La categorization sociale. In Mscovici, S., editor, *Introduction a la psychologie sociale*, volume 1, pages 272–302.

Tang, J., Yankelovich, N., Begole, B., Van Kleek, M., Li, F., and Bhalodia, J. (2001). Connexus to awarenes: Extending awareness to mobile users. *Proceedings of SIGCHI*.

Taube, V. (2004). Connected and disconnected? on the impact of internet use on social connectedness. *Computational and Mathematical Organization Theory*, 10(3):227–241.

Taylor, A. (2002). Teenage 'phone-talk' and its implications for design. *Proceedings of NordiCHI 2002, Bridging the gap between field studies and design workshop*.

Thórisson, K. (1997). Gandalf: An embodied humanoid capable of real-time multimodal dialogue with people. *Proceedings of the First ACM International Conference on Autonomous Agents*.

Thorndike, E. (1920). Intelligence and its uses. *Harpers Magazine*, (140):227–235.

Tory, M. and Staub-French, S. (2008). Qualitative analysis of visualization: a building design field study. *Proceedings of the 2008 Conference on Beyond Time and Errors: Novel Evaluation Methods For information Visualization*.

Van Breemen, A. (2004). Bringing robots to life: Applying principles of animation to robots. *Proceedings of the CHI2004 workshop Shaping Human-Robot Interaction*.

Van Doorn, M. (2009). *End-user Programming of Ambient Narratives for Smart Retail Environments*. Eindhoven University of Technology.

Van Doorn, M., Van Loenen, E., and De Vries, A. (2008). Deconstructing ambient intelligence into ambient narratives: the intelligent shop window. In *Ambi-Sys '08: Proceedings of the 1st international conference on Ambient media and systems*, pages 1–8, ICST, Brussels, Belgium, Belgium. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

Venkatesh, V., Morris, M., Davis, G., and Davis, F. (2003). User acceptance of information technology: Towards a unified view. *MIS Quaterly*, 27(3):425–478.

Vernon, P. (1933). Some characteristics of the good judge of personality. *Journal of Social Psychology*, (4):42–57.

Vetere, F., Gibbs, M., Kjeldskov, J., Howard, S., Mueller, F., Pedell, S., Mecoles, K., and Bunyan, M. (2005). Mediating intimacy: designing technologies to support strong-tie relationships. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '05, pages 471–480, New York, NY, USA. ACM.

Weiser, M. (1991). The computer for the 21st century. *Scientific American*, 265(3):94–104.

Weiss, R. (1973). *Loneliness: The experience of emotional and social isolation*. Cambridge, MA:MIT Press.

Westbrook, R. (1980). Intrapersonal affective influences upon consumer satisfaction. *Journal of Consumer Research*, (7):49–54.

Wirtz, J. and Bateson, J. (1995). An experimental investigation of halo effects in satisfaction measures of service attributes. *International Journal of Service Industry Management*, 6(3):84–102.

Wright, D., Gutwirth, S., Friedewald, M., Vildjiounaite, E., and Punie, Y. (2008). *Safeguards in a World of Ambient Intelligence*. Springer.

Wu, B. and Petroshius, S. (1987). The halo effect in store image management. *Journal of the Academy of Marketing Science*, 15:44–51.

Zhao, Q. and Stasko, J. (1998). The awareness-privacy tradeoff in video supported informal awareness: A study of image-filtering based techniques. *Technical Report GIT-GVU-98-16*.

Zhao, Q. and Stasko, J. (2002). What's happening?: promoting community awareness through opportunistic, peripheral interfaces. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '02, pages 69–74, New York, NY, USA. ACM.

Appendix

Appendix A: Pleasure - Dominance - Arousal Scale

Ratings are on a 9 point scale ranging from -4 to +4

(A)	WIDE AWAKE										SLEEPY
(D)	CONTROLLED										CONTROLLING
(P)	MELANCHOLIC										CONTENT
(A)	AROUSED										UNAROUSED
(D)	INFLUENTIAL										INFLUENCED
(D)	AWED										IMPORTANT
(P)	SATISFIED										UNSATISFIED
(D)	AUTONOMOUS										GUIDED
(P)	BORED										RELAXED
(P)	HAPPY										UNHAPPY
(D)	DOMINANT										SUBMISSIVE
(A)	CALM										EXCITED
(P)	ANNOYED										PLEASED
(P)	HOPEFUL										DESPAIRING
(D)	CARED FOR										IN CONTROL

(**A**)	RELAXED	□□□□□□□□□	STIMULATED
(**A**)	DULL	□□□□□□□□□	JITTERY
(**A**)	FRENZIED	□□□□□□□□□	SLUGGISH

Appendix B: Social Behaviors Questionnaire

Ratings are on a 5-point scale: agree, slightly agree, do not agree or disagree, slightly disagree, disagree.

The robotic cat:
... is skilled in handling social situations.
... does not care about others' needs.
... has a high opinion of itself.
... likes to watch others and study them.
... makes itself the center of attention.
... tries to lead others.
... makes enemies.
... seeks to influence others.
... suddenly loses interest.
... handles tasks smoothly.
... thinks highly of itself.
... feels the emotions of the other.
... takes others' interests into account.
... does not see the consequences of things.
... takes time out for others.
... trusts others.
... likes to anticipate what the a person will do at a given time.
... boasts about its virtues.
... takes no time for others.
... is able to cooperate with others.
... is hard to understand.
... believes others should take care of themselves.
... is preoccupied with itself.
... gets bored easily.
... takes control of things.
... trusts what others say.
... likes to be able to help others.
... thinks about the disposition of others.
... values cooperation over competition.
... acknowledges others accomplishments.
... misrepresents the facts.
... knows how to get things done.
... tells the truth.
... anticipates the needs of others.
... is indifferent to the feelings of others.

... dislikes talking about itself.
... reassures others.
... is concerned about others.
... is not interested in others' problems.
... comes up with good solutions.
... does not understand things.
... lives in a world of its own.
... breaks its promises.
... waits for others to lead the way.
... does the opposite of what is asked.
... makes others feel at ease.
... is polite to strangers.
... believes that others have good intentions.
... breaks rules.
... believes in human goodness.
... makes others feel good.
... believes that it is better than others.
... loves to help others.
... seems interested in what goes on in others' minds.
... completes tasks successfully.
... keeps its promises.
... says inappropriate things.

Appendix C: User Satisfaction Questionnaire

Ratings are on a 5-point scale: disagree, disagree, do not agree or disagree, slightly sagree, agree.

Overall impression
(1) I like the way in which things are indicated on this APP
(2) Things are badly indicated on this APP
(3) Operating this APP is pleasant
(4) Operating this APP is difficult
(5) I am content with this APP
(6) It is clear to me what this APP is doing when it is busy
(7) It is frustrating to use this APP
(8) The appearance of this APP makes it easy to operate
(9) I think this APP is comple
(10) This APP has the possibilities I expected based on its appearance
Efficiency
(11) The workings of this APP are such that it is easy for me to operate
(12) I have to do a lot of things to operate this APP
(13) Operating this APP takes effort
(14) Operating this APP is made easy by the design of the buttons, symbols and text
(15) The wording and symbols on this APP make clear which elements are important to achieve my goals
(16) I can achieve my goals without difficulty
(17) The way this APP works gets in my way when trying to achieve my goals
(18) I am satisfied about the ease with which I can operate this APP
(19) It is difficult to find the information that I need
(20) It takes effort to correct unwanted outcomes
(21) I can easily correct unwanted outcomes
(22) It is easy to do the things I intend to do with this APP Learnability
(23) The workings of this APP make it easy to learn how to operate it
(24) The design of the buttons, symbols and text makes it easy to learn how to operate this APP

(25) It is difficult to remember the meaning of the wording and symbols of this APP
(26) It is hard to learn how to operate this APP
(27) It is easy to learn how to operate this APP
(28) I am satisfied with the ease with which I can learn how to operate this APP
(29) The reactions this APP gives to inform me about its current state help me to learn how to operate it
Transparency
(30) It is difficult to know whether everything is okay
(31) The information provided by this APP is organised in a clear way
(32) I understand the wording and the symbols on this APP
(33) The wording and the symbols on this APP are incomprehensible
(34) When I want to reach a goal, information is available on how far I am and what I have to do
(35) The relation between what I do and what happens is unclear
(36) I know in advance what the effect of an action will be
(37) The reactions of this APP on my actions are incomprehensible
(38) It is clear to me how I can achieve my goals
(39) The information provided by this APP is incomprehensible
(40) The reactions of this APP are incomprehensible
Conformity
(41) The design of the buttons, symbols and text differs from what I would expect for this type of product
(42) The appearance of this APP differs from what I would expect for this type of product
(43) I would expect other wording and symbols on this APP than what it shows
(44) The wording and symbols on this APP are familiar to me
(45) This APP works according to my expectations
(46) This APP has the possibilities I would expect from this type of product
Enjoyability
(47) This APP looks nice
(48) This APP is ugly
(49) The possibilities of this APP make it interesting to have

(50) I am pleased with the possibilities of this APP
(51) Using this APP is an unpleasant experience
(52) I enjoyed using this APP
(53) I would like to have this APP
(54) Using this APP is boring

List of Abbreviations

AAL	Ambient Assisted Living
AmI	Ambient Intelligence
AmI	Ambient Intelligence
BMIS	Brief Mood Introspection Scale
EARC	Experience and Application Research Centers
EARC	Experience and Application Research Centers
GAS	Group Attitude Scale
ICT	Information and Communication Technologies
IPIP	International Personality Item Pool
IPO-SPQ	Instituut voor Perceptie Onderzoek - Social Presence Questionnaire
LON	Local Operated Network
NEO-FFI	Neuroticism-Extroversion-Openness Five Factor Inventory
NEO-PI	Neuroticism-Extroversion-Openness Personality Inventory
PDA	Pleasure, Dominance and Arousal scale
SBQ	Social Behaviors Questionnaire
TLX	NASA Task Load Index
UCD	User - Centred Design
USQ	User Satisfaction Questionnaire
USQ	User Satisfaction Questionnaire
UTAUT	Unified Theory of Acceptance and Use of Technology

Social Interactions in Ambient Intelligent Environments

This thesis is focussing on research into Ambient Intelligence. More specific, the thesis will focus on *infrastructures*, *methodology* and *applications* deployed in Ambient Intelligence research. By presenting work on those three aspects the thesis contributes to the scientific research field of Ambient Intelligence research. It should be emphasized that this does not imply that all Ambient Inteligence related research should be done within the presented infrastructures and with the presented methodology.

The applications of Ambient Intelligence presented in this thesis investigate the possible role of technology to support people in maintaining social connectedness. More specific, a series of empirical studies is reported that explore the potential for *Ambient Intelligent* environments to support and facilitate social affiliation through technology. By positioning the human need in the center of technology development, Ambient Intelligence (described in chapter 2) is an attractive approach to application development for social affiliation. The studies reported in this thesis have in large part been conducted in the ExperienceLab at Philips Research. These test facilities (described in chapter 4) are specifically suitable for experience research (described and illustrated with a case study in chapter 3) that requires controlled yet realistic testing environments. In a first series of experiments with *awareness systems* to mediate presence through silhouette representations of remote activities is investigated (see chapter 5) and mediating availability through presence information (see chapter 6). The results of these studies show that providing simple presence information can support feelings of social connectedness while shielding for privacy. Additional, these awareness systems support the function of availability judgements (i.e. using the presence information to estimate the availability of the remote person to engage into a conversation).

Taking Ambient Intelligence (AmI) to the next level, the role of *social intelligence* (see the extended AmI model in chapter 2) is investigated in a controlled laboratory study and discussed in chapter 7. From this study it is learned that equipping technical systems with a level of social intelligence results in enriched and human-like interactions with the system.

The scientific contribution of this thesis is in:

- ◇ A Methodological Framework for Ambient Intelligent research
- ◇ The empirical validation of Experience and Application Research Centers as scientific instruments
- ◇ Demonstration of the affective benefits of Awareness Systems
- ◇ Illustration of how social intelligence can be manifested in a Home Dialogue System
- ◇ The Assessment of Social Intelligence in a Home Dialogue System
- ◇ Illustrated the spillover effects of Social Intelligence

Acknowledgments

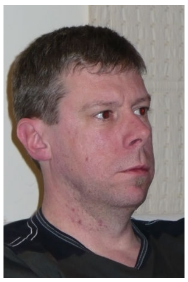
Many of the studies reported in this thesis have been carried out in cooperation with students from the TECHNICAL UNIVERSITY EINDHOVEN, the TECHNICAL UNIVERSITY DELFT and the UNIVERSITY OF TILBURG. I would like to acknowledge specifically Claire Huijnen, Rick van Geel, Elisabeth Leegwater, Ehsan Baha and Privender Saini for the cooperation in the reported studies. A special acknowledgment to Prof. Dr. Panos Markopoulos for his support and supervision in the work that has led to this thesis. Our cooperation has crossed boundaries between organisations and geographical locations. I am looking forward to continuing this fruitful cooperation in the years to come.

The most important person in this challenging trajectory has been undoutably Prof. Dr. Emile Aarts. Not only for being my promotor in this work but more specific for providing a continuous drive and motivation to achieve this result. Without his support this thesis would never have been finalized. I would like to express my sincere gratitude towards the members of the reading committee Prof. Dr. Berry Eggen, Prof. Dr. Huib de Ridder and Prof. Dr. Anton Nijholt for their efforts in reviewing this thesis. I would like to thank Prof. Dr. Manfred Tscheligi and Dr. Dr. Norbert Streitz for their role as part of the corona for the PhD defense.

My dear colleagues Elly Pelgrim-Zwartkruis and Privender Saini: thank you for accepting the burden to be my paranymphs during the ceremony. Your presence and support means a lot to me.

Albeit cliché, I would like to thank my family and in specific my Wife and children. In the last couple of years they have accepted my absence during evenings, weekends and even vacations to work on this thesis. I thank them for giving me the opportunity to work on this thesis.

Curriculum



Boris de Ruyter was born on 11 November 1966 in Kortrijk, Belgium. After my graduation as an experimental psychologist at the University of Ghent in Belgium, I completed my military service at the Military hospital in Brussels after which I worked as a research assistant at the University of Antwerp. In 1994 I joined Philips Research in Eindhoven where I am appointed as Principal Scientist.

With more than 63 publications (conference proceedings, journal articles and book chapters) and several patents, I have been involved in User Centered research in the domain of Health & Wellbeing applications at Philips Research. Some of this research I have conducted in the context of EU funded research projects in which I held different coordinative roles. Besides my role as Principal Scientists at Philips Research, I have been involved in several academic events (SIGCHI, AVI, AMI and MobileHCI) in different roles (co-chair, program chair, track chair, program committee member).