UNIVERSITY OF TWENTE.

SERA

Deliverable 3.1

"Design Guidelines"

Contract number: **FP7-231868 SERA**

Social Engagement with Robots and Agents

The research leading to these results has received funding from the European Community's Seventh Framework Programme (*FP7/2007-2013*) under *grant agreement* n° 231868.

European Commission

SEVENTH FRAMEWORK PROGRAMME

Identification sheet

Project ref. no.	**FP7-231868**
Project acronym	SERA
Status & version	Version "2.0"
Contractual date of delivery	January 31, 2010
Actual date of delivery	February 17, 2010
Deliverable number	D3.1
Deliverable title	Design Guidelines
Nature	Report
Dissemination level	**PU** Public
WP contributing to the deliverable	WP3
WP / Task responsible	University of Twente
Editor	Dirk Heylen
Editor address	Human Media Interaction University of Twente POBOX 217 7500AE Enschede The Netherlands
Author(s) (alphabetically)	Somaya Ben-Allouch, Betsy van Dijk, Sabrina Eimler, Dirk Heylen, Iwan de Kok, Tineke Klamer, Nicole Krämer, Sabine Payr, Paolo Petta, Astrid von der Pütten, Ronald Poppe, Bart van Straalen, Mariët Theune
EC Project Officer	Pierre Paul Sondag
Keywords	Design guidelines, literature review, robot appearance, communicative skills, social skills
Abstract (for dissemination)	This deliverable considers two questions: what would social agents and robots have to look like and what skills should they have. We probe these aspects by looking at the recommendations and experience laid down in the literature and by conclusions drawn from the preliminary data analysis of the first iteration. Rather than presenting a full survey of the literature, we discuss what we believe to be a relevant sample that is sufficient for current purposes to identify the range of dimensions to consider in designing social agents or robots.

Contents

1 Introduction 4
2 The Literature Sample 4
3 Guidelines from the Literature 5
3.1 Appearance 6
3.2 Communicative skills 10
3.3 Social Skills 13
4 Summary and Conclusions 16
References 18

1 Introduction

Among the deliverables for WP3 are two reports, one on Design Guidelines and one on Reference Architecture. The design guidelines and reference architecture documents are concerned with questions related to three aspects of social agents or robots: functionality, appearance and components.

- Functionality: what should they be able to do, and related questions
- Appearance: what should they look like, and related questions
- Components: how should they be build, and related questions

A discussion of the first two aspects is presented in this deliverable: D3.1. The third aspect is dealt with in Deliverable D3.2 (Reference Architecture). We will probe these aspects by looking at the recommendations and experience laid down in the literature. Rather than presenting a full survey of the literature, we will discuss what we believe to be a relevant sample that is sufficient for current purposes to identify the range of dimensions to consider in designing social agents or robots. Besides identifying these dimensions, we will highlight the diversity of grounds on which recommendations have been made as the research on these issues is still very much work in progress. Some recommendations are based on conjectures whereas others derive from smaller or larger scale empirical studies. Although empirical studies may seem to provide a firmer basis for guidance, the findings may apply specifically to the domain, application, system or user group from the study. This should be taken into account in the use of the guidelines in other work packages. The purpose of this document is to identify the major factors to guide the design of a social agent or robot based, primarily, on the experience and recommendations found in the literature.

This document is intended to be of use in three work packages. First, it may provide dimensions for WP1 (Theory and Analysis) to consider in analysing the data obtained in the various iterations. Second, it provides insights from the evaluation of systems that resemble the ones that are being developed in WP2 (Field Studies) and WP4 (Showcase) which might be used to improve these systems, or guide their design. Third, it provides motivation for central components to consider in the reference architecture.

We have sampled from the literature a number of studies regarding social agents and robots and looked for statements about what such a system "*should look like*", what it "*should be able to do*", what users "*prefer*", or what "*works best*". As we will discuss below, this does result neither in a predefined design nor in a golden standard of guidelines that social agents and robots should adhere to but rather it gives us a list of recurrent concerns, along with some methods and examples that can guide the design process.

In the next section we discuss the nature of the studies that we have sampled and analysed. Section 3 presents the requirements that have been considered in the sample literature. The concluding Section 4 provides a first confrontation of guidelines and the data from the first field study iteration (WP2).

2 The Literature Sample

The past decade has shown a clear growth in interest to study long-term relationships between humans and artificial companions such as agents and robots, witness the current surge in projects such as Companions, LIREC, and SERA. Tim Bickmore's PhD study on the agent Laura (Bickmore, 2002), was one of the first longitudinal studies in which users interacted with an agent for longer periods of time in their homes, rather than the one-off short laboratory session. These kinds of longitudinal studies are still exceptional. Most empirical studies on human-robot interaction describe experiments that were performed in controlled laboratory settings.

Besides the longitudinal environmental and the shorter laboratory experiments, there are also more general studies on human-computer or human-human interaction from which guidelines can be derived. The extrapolation of findings from human-human studies to human-agent interactions can be based on the idea that humans have a particular way of interacting that comes natural to them so that a system that exploits this type of interaction is also easier to learn and experienced as more pleasant. It can also be based on the goal of building agents or robots that are as much human-like as possible. The validity of extrapolating guidelines for the design of naturally interacting artificial systems from what we know of human-human interactions is often argued for by the Media Equation, which was demonstrated in a series of studies by Reeves and Nass and students (Reeves and Nass, 1996). Many of these first studies just used "old-fashioned" text-based interfaces, but since then evidence has been steadily extended in support of the broader claim that in interaction with computers, humans show the same relational stance towards the computer as they show towards humans. For instance, studies on this so-called CASA ("Computers Are Social Actors") paradigm have shown that computers that use flattery, or which praise rather than criticize their users are liked better (see also Fogg and Nass, 1997; Johnson et al., 2004). Also, users prefer computers that match them in personality (Nass and Lee, 2001, Johnson and Gardner, 2007), a phenomenon known as the "similarity attraction" principle in social psychology (Byrne et al., 1986). In the deliverable D1.2a, which mainly reviewed social psychological theories on human-human interaction, several guidelines for human-computer interaction were extrapolated in this way from the psychological literature. The deliverable at hand can be seen as complementing D1.2a, by focusing on systems rather than theory.

As mentioned above, this deliverable is based on the analysis of a sample of the literature on social agents and robots. The selection is biased towards studies on systems that are used for healthcare or health promoting actions, though also studies that are more general or that focus on robots as companions are taken into consideration. The following roles for agents and robots are found in the literature considered: social companion, exercise counselor, chess player, health adviser (weight loss, diet, exercise), reminder service. Many studies were concerned with either older adults or children.

3 Guidelines from the Literature

Several concerns appear over and over again in the literature. Fong et al. (2003) listed the following six types of requirements which are often cited. Social robots and agents should be able to:

1. Express and/or perceive emotions
2. Communicate with high-level dialogue
3. Learn/recognize models of other agents
4. Use natural cues (gaze, gestures, etc.)
5. Exhibit distinctive personality and character
6. If possible, learn/develop social competencies

We refine the guidelines from the literature into the nine classes below, splitting up some items of the above list and adding others found in the literature. The first class that we will cover deals with appearance. The other classes deal with behaviours and skills, i.e. the functional level. Classes 2 to 5 deal with the more basic behavioural, communicative skills. Classes 6 to 9 are actions on a higher level related to the social skills that an agent or robot should have.

1. Appearance and embodiment
2. Range of conversational behaviours
3. Use of natural or non-natural cues to use in interaction

4. Display of cognitive and affective signals
5. Need to recognize affect
6. Relational social skills involved in communication such as meta-relational communication, politeness, facework, relationship repair, or the ability to draw users into repeated interactions
7. Outcome of the relational skills on social psychological variables such as empathy, working alliance, trust, liking, engagement
8. Need for personalisation, adaptation, and tailoring to user or task
9. Notion of an autonomous, character with memory and personality

Section 3.1 will look at **appearance** (1), Section 3.2 at **communicative skills** (2-5) and Section 3.3 at social skills and personality, or the **higher level skills** (6-9). This division thus more or less corresponds to a distinction of hierarchical levels: social skills and personality are implemented using the communicative skills and in order to communicate and be expressive the robot or agent should be embodied in specific ways.

3.1 Appearance

While in Deliverable 1.1 appearance was already mentioned to be an important aspect, the question that remains to be answered is what exactly the robot/agent should look like. Generally it is argued that it is desirable to design "*life-like*" agents. Several reasons may be given. First, it is often argued that agents should resemble people because they perform roles and fulfil tasks that are normally performed by humans. Second, users are used to and prefer to interact with other humans. On the other hand, it has been claimed that social robots should not resemble humans too much. When robots look like humans this may elicit strong expectations about the robot's social and cognitive competence. If such expectations are not met, then the user is likely to experience confusion, frustration and disappointment (Dautenhahn, 2002). Dautenhahn (2002) argues that this effect is highly context dependent, though. In some situations it is more acceptable that expectations are not met than in other situations. For example, when a robot serves as a servant for elderly people it is less acceptable when expectations are not met than in situations where robots serve as a kind of toy (e.g. AIBO).

In this section we first consider some studies that deal with the question of embodiment: should the system be embodied as a robot, as an agent, or does it suffice to have something like a disembodied text-based interface? Next, we consider some studies that look at the appearance of robots and agents, e.g., whether female-looking robots are preferred over male-looking robots. After that we present some studies dealing with the factor "task" and how this influences the preference for appearance, and we conclude with a note on the "uncanny valley".

Types of Embodiment

Several studies were conducted to determine whether a robot or agent interface would be preferred over other types of interfaces such as a simple text interface, for instance. Looije et al. (2006) and Looije et al. (to appear) showed that both children and older adults preferred to interact with an animated iCat robot over a simple text interface in studies on support of diabetes treatment. A pilot study by Tapus et al. (2009) with a socially assistive robot for elderly people with cognitive impairments showed that the robot's physical embodiment played a key role in its effectiveness. They compared an embodied system with a computer interface in the context of a music-based cognitive game. Not only did users prefer the embodied system, but the embodied system also had an overall effect on sustaining and improving the performance on the task.

A comparison between a robot and a virtual agent helping with the Tower of Hanoi puzzle was carried out by Wainer et al. (2006). In this work, it was shown that physical embodiment has a measurable effect on performance and perception of social interactions. They also compared the condition of a robot which was physically present to one that was tele-present. The co-located physical robot was perceived to be more watchful. It was also more enjoyable than either the virtual agent or the remote robot.

These results seem to suggest that robots are generally preferred above agents and agents are preferred above text interfaces. But the novelty effect may be an important factor in these studies and it is not clear how soon it would wear off. Kidd (2008) points out that the appearance of a robot seems to be a factor for a minority of people only, both in the case of initial attraction and in longer-term engagements. More important for people is what the robot has to offer and whether it is useful.

Guidelines:

- **Prefer robots over agents and agents over non-embodied interfaces. However, take into account the novelty effect when designing your robot/agent.**
- **Overall, functionality is more important than appearance.**

Appearance of Robots

An important theme of research regarding humanoid robots is appearance (Goetz et al. 2003; Robins et al. 2004; Koay et al. 2007; 2009; Nomura et al. 2007; Walters et al. 2008). Walters et al. (2008, p.164) introduce the following terminology for robot appearances: a *mechanoid* robot is relatively machine-like; a *humanoid* robot is not realistically human-like and readily perceived as a robot by human interactants (e.g. wheels for locomotion), but possesses some human-like features which are often stylised (e.g., facial features); an *android* robot exhibits appearance (and behaviour) as close to human as technically possible.

Goetz et al. (2003) tested two hypotheses, the positivity hypothesis and the matching hypothesis using the robot Pearl. It was concluded that a robot's appearance and behaviour influences people's perceptions on the robot and their willingness to listen to the robot's instructions. Robins et al. (2004) tested two versions of Robota (a 45cm high robotic doll) with four autistic children (age 5-10): one dressed, "pretty girl" version and one undressed, mechanoid version. The experiments were long-term, 13 trials on average per child over a period of several months. The results showed that the autistic children preferred interaction with the robot in its plain robotic appearance over the 'pretty girl' appearance, but over time they became familiar with both appearances of the robot. There was also an indication that the plain appearance was a salient feature in causing the children to use the robot as a mediator to interact with adults around them (investigators and carers): *"In some cases the children started to use the robot as a mediator, an object of shared attention. They opened themselves up to include the investigator in their world, actively seeking to share their experience with him as well as with their carer."* (Robins et al. (2004) p280). Koay et al. (2007; 2009) tested four different robot appearances (short and tall, humanoid and mechanoid) over five weeks. The results showed that participants' preferences changed over time as the participants tended to habituate to the robot. This trend was significant in terms of appearance and approach direction. The results seemed to indicate that participants who were accustomed to the robot preferred the more 'in control' situation. When unexpected events occurred, participants appreciated reduced robot autonomy. Walters et al. (2008) studied robot appearances as well. They did an experiment with three different kinds of robots: a mechanoid robot, a basic (humanoid) robot and a humanoid robot[1] – android robots were deliberately excluded given results from previous surveys (Khan, 1998; Scopelliti et

[1] In contrast to the humanoid robot equipped with two seven-degrees of freedom arms capable of human-like gestures (e.g. waving), the basic robot was equipped with just one single-degree of freedom arm which enabled it to make pointing gestures. The mechanoid robot was equipped with a simple single-degree of freedom gripper that could only be moved up or down.

al, 2004; Dautenhahn et al., 2005; see also the introduction to the present subsection). Students did not interact with the robots themselves, but expressed their opinions about the robots in a video-based human-robot interaction trail. The results showed that overall participants tended to prefer robots with more human-like appearances and attributes; however it was also found that these preferences seemed to differ with the students' personalities: introverts and participants with lower emotional stability tended to prefer the mechanical appearance to a greater degree than other participants.

Guidelines

- **No absolute preference can be given to the appearance of robots in terms of humanoid vs. mechanoid. Make the choice dependent at least on preferences in the user group, as personality and other user characteristics play a role.**
- **Take into account habituation processes in piloting appearance.**

Appearance matched to task

Goetz et al. (2003) tested two contrasting social psychology hypotheses using the humanoid Nursebot robot Pearl[2]: The "positivity hypothesis" asserts that attractiveness, extraversion and cheerfulness correlate generally with acceptance and compliance by human peers, while the "matching hypothesis" states that appearance and social behaviour of a robot should be matched to the seriousness of the task and situation. In the study people were presented with pictures of robots that differed along three dimensions: age, sex, and human versus mechanic looking. Participants were asked which robot appearance they would prefer for different tasks. For female-looking robots, participants preferred the humanlike robots to the machine-like robots for most jobs: actress and drawing instructor (artistic jobs), retail clerk and sales representative (enterprise jobs), office clerk and hospital message and food carrier (conventional jobs), aerobics instructor and museum tour guide (social jobs). Machinelike robots were preferred over the humanlike robots for jobs including lab assistant and customs inspector (Investigative) and for soldier and security guard (realistic). The patterns for the masculine looking robots were not as strong but went generally in the same direction. In this case participants slightly preferred humanlike robots for artistic and social job types, but preferred machinelike robots for realistic and conventional jobs. It was concluded that a robot's appearance and behaviour indeed do influence people's perceptions of the robot's capabilities and social behaviour skills and their willingness to listen to the robot's instructions. Furthermore, elicited responses were framed by people's expectations of the robot's role in the situation. Thus, it was the more differentiated matching hypothesis that found support, e.g. in terms of more humanlike, attractive, or playful robots not necessarily being considered to be more compelling, in particular when such appearance did not match the given task context.

A similar experiment was carried out by Hegel et al. (2009). They looked at the influence of visual appearance of social robots on judgments about their potential applications. They chose videos of twelve robots, Barthoc Jr., iCat, AIBO, BIRON, KeepOn, Kismet, Leonardo, Robovie, Repliee Q2, ASIMO, Paro, and Pearl and thirteen "jobs": security, research, healthcare, personal assistant, toy, business, pet, entertainment, teaching, transport, companion, caregiver, public assistant. Among the results of the study was that the application fields entertainment, pet, toy, and companion strongly correlate with an animal-like appearance, while the others strongly correlate with human-like and functional-like appearances. The visual appearance of robots was shown to be a significant predictor for the estimation of applications. Additionally, the ratings showed an attractiveness bias, in that robots that were judged as more attractive were also evaluated more positively on a liking scale.

[2] http://www.cs.cmu.edu/~nursebot/ (visited: 2010-02-11)

Guidelines

- **Consider whether appearance and role/task should be matched. (Note that the studies mentioned did not involve actual interactions with robots or agents, so novelty effects and habituation may influence the importance of this guideline over time)**
- **Pay attention to making the robot/agent attractive.**

The Uncanny Valley

Masahiro Mori tried to predict the psychological effects that different robotic designs and other human-like artefacts have on humans (Dautenhahn, 2002). Mori proposed the uncanny valley theory, in which he predicted that the more life-like we make robots, the more believable they become, with a maximum believability being reached in the case of full similarity with human beings. However, the transition towards full human appearance incurs a local minimum, characterised by a sharp drop in familiarity when robots start to appear so life-like they might be taken for " 'real' " (Dautenhahn, 2002). In this case, robots can cause a sudden uncanny and unpleasant feeling, because still existing differences suddenly make us realise that the robots are not human after all and fail to match our expectations. In addition, Mori distinguished two separate graphs, reflecting two different components of similarity to human beings, namely movement and appearance. The movement curve is considered to be more dominant than the appearance curve. Thus, people feel more uncanny when there is something unfamiliar about the way a robot moves than about the way the robot is looking. Duffy (2003) concludes that with respect to appearance, mechanic, more iconic heads would be preferred above those that attempt to resemble human heads.

Studies that tested the uncanny valley hypothesis concerning robots include the following:

- The results of the study by Groom (2009) support the uncanny valley hypothesis: the agent in the mixed-realism condition received the most positive ratings, the most realistic agent was rated the lowest. Overall, participants' negative response to the most realistic agent was as clear as their preference for agents that demonstrated inconsistent, moderately realistic behaviour.
- The results of MacDorman (2006) do not indicate a single uncanny valley for a particular range of human likeness. Along the lines of Mori's more detailed analysis, the results suggest that human likeness is only one of perhaps many factors influencing the extent to which a robot is perceived as being strange, familiar, or eerie.
- The results of Oyedele (2007) support the hypotheses that individual responses to the images shown would resemble a pattern similar to the uncanny valley and that both the context within which humans and robots interact and the extent to which robots resemble human beings influence the experienced affect towards the robot.
- The results of Walters et al. (2008) show that participants tended to prefer robots with more human-like appearances and attributes. However, introvert participants and participants with lower emotional stability tended to prefer the mechanical looking appearance to a greater degree than other participants. The results also show that it is possible to rate individual elements of a particular robot's behaviour and then assess the contribution (or lack thereof) of that element to the overall perception of the robot by people. Relating participant's ratings of individual robots to independent static appearance ratings provided evidence that supports Mori's theory of the uncanny valley.
- Kanda et al. (2008) compared two humanoid robots, ASIMO and Robovie, and a human according to appearance in simple interactions at first meeting. We summarise their findings: 1. Subjective impressions: ASIMO received better subjective impressions than Robovie or the human; 2. Verbal responses: participants gave the same amount of information with identical politeness to ASIMO, Robovie, and the human; 3. Distance at first conversation:

participants tended to stand closer to ASIMO than to Robovie or the human; 4. Degree of participant waist angles when bowing as a greeting: participants bowed more deeply to the human than to Robovie; 5. Delay time of vocal response to greetings: participants replied more rapidly to humans than to the robots, and they replied more rapidly to ASIMO than to Robovie; 6. Extent of each arm's movements during talking: no significant difference was found; 7. Participant delay time of gaze response to pointing: participants looked most rapidly at a poster when Robovie pointed to it; 8. Distance while walking: participants tended to walk closer to ASIMO than to Robovie or the human; 9. Speed during walking: participants walked faster with the human, but this largely reflected the relative speeds of the robots and the human.

Guideline

- **To test different robot/agents designs with respect to uncanny valley effects, consider at least the following factors: (1) the variables related to the location of the different designs along the scale ranging from unequivocal artefact appearance to human resemblance and (2) the effect is dependent on the type of user.**

Summary

From this short overview one can draw a few conclusions. First, studies that compared anthropomorphic interfaces with simple text-based interfaces found that users had a preference for the former kinds. Second, there are several variables that should influence the design or choice of robot appearance: in addition the task at hand, personal preferences are also of importance.

In general, one can conclude that carefully set-up empirical (pilot) studies can inform the definition of the appearance of an agent or robot so that it is matched with the task and the user group. Habituation processes may still make up for less than optimal choices. Kidd (2008) provides an important point: functionality is more important than appearance.

3.2 Communicative skills

The ability to communicate, verbally and/or non-verbally, is a common-sense requirement for artificial companions or social robots. Deliverable 1.1 already refers to the fact that theories about verbal and nonverbal behaviour might be of relevance.

Conversational behaviours

Some form of communication is needed as a basis to establish a social relation (Castellano et al. (2008); Duffy (2003), Fong et al. (2003); Green et al. (2004), Kidd (2008); Li et al. (2006)). In embodied conversational agents this is often interpreted as a requirement for social robots and agents to possess conversational skills similar to humans, including ways to open and close conversations and showing engagement (Sidner et al. 2004); turn-taking; providing feedback (Kidd, 2008), contrast, and emphasis; showing attention; addressing (Bruce 2002); formulating sentences, and constructing multimodal communicative actions (see Cassell et al. 1999 for an overview). On the basis of such communication skills, further social interaction skills can be built. For instance, providing feedback can be important to motivate users or to express empathy (Blanson-Henkemans et al. 2009; see also below). Skills like these are often assumed to be important in human-robot interaction just because they are fundamental in human-human interaction. However, they apply in particular to humanoid robots and agents with dialogue abilities. For the case of zoomorphic agents/robots the case might be different (consider, for instance, the examples of Paro or AIBO). A tentative conclusion is given in the following guideline:

Guideline

- **If one equips robots and agents with conversational skills, one needs to take into account the many levels on which conversation works and consider how to contribute to verifying, establishing and maintaining required preconditions and context, accounting for procedures such as of engagement, feedback, and information structuring.**

Use of natural cues

Several authors focus on the importance of using “natural” cues in interaction – the verbal and nonverbal cues humans use effortlessly in interaction. They have studied the use of particular mainly nonverbal behaviours and their effects on the interaction and the social relationship: eye contact, look-at behaviours, head, arm, and hand gestures (see e.g. the references in the previous paragraph). Subtle gestures were found to have a positive effect on understanding what a robot was doing (Breazeal et al. 2005). One of the factors that Kidd (2008, p. 162) mentions as contributing significantly to the creation of a successful system is the appearance of eye contact and the movement of the eyes: “Numerous participants who had the robot (and countless others who have seen or interacted with it for short durations) noted during the final interview that the eyes of the robot and the fact that it looked at them drew them into interactions with the system and made it feel more lifelike. Users were clearly more engaged with a system that looked at them […].” Also Kozima et al. (2003) point out the role of eye contact on joint attention as a prerequisite for social interaction. Bruce et al. (2002) indicate attention of the robot by having the robot turn to a person when it wants to address that person (see also the work by Sidner et al.(2004, 2005) on engagement mentioned above).

Interestingly, some conversational functions can also be expressed by cues that are less natural. Kobayashi et al. (2008) used a blinking LED to notify a user about the robot’s internal state such as processing or being busy. The use of such a visual device made conversations go smoother, human users would not repeat themselves as often (see also Sengers 1999).

Guidelines

- **As humans will interpret the (nonverbal and other) behaviours of a robot/agent similarly to the behaviours of humans, the behaviours should be carefully designed and work in similar ways as in the human case (e.g., provide non-functional clues) to avoid misinterpretation and to benefit from the inferences made.**
- **On the other hand, robots and agents are accepted to be different from humans and can have non-human-like interfaces that can be exploited in the interaction to take care of important communicative functionalities.**

Affective and social behaviours

Kobyashi et al. (2008)’s blinking LED highlights the importance for robots and agents to express internal state. Many authors emphasize the importance of the capability and ability to express social, affective behaviour (Sengers 1999, Bartneck 2003; Bickmore 2003; Blanson-Henkemans et al. 2009; Bruce et al. 2002; Cassell et al. 1999; Duffy 2003; Fong et al. 2003; Kessens et al. 2009, Kidd 2008, Castellano 2008 and 2009; Looije et al. 2009). According to Kidd (2008) expressing the internal state to the user is a most critical factor for a robot to be accepted and used. Castellano et al. (2009) claim that companion robots should be able to display and recognize affective behaviour in order to sustain long-term interactions with users. No further motivation is given, however. In another paper (Castellano et al 2008) this assertion is generalized to other mental states.

With respect to expressivity, the behaviours displayed by humans in the human-human setting are often taken as the primary model. Bickmore (2002) noticed in his study on human-human interactions (between trainer and client) that the trainer mainly used three expressions that were important to mark affective state or context: a neutral face (majority of cases and in particular during information exchanges), a concerned face (during empathy exchanges) and a happy/smiling face during ritual actions (greetings and farewells), social dialogue and humour. That different task contexts may require different expressions was also pointed out by Kessens et al. (2009) who looked at the role of emotional expressions (face and voice) for different roles of a health-promoting robot (the iCat was used). The iCat would act as a companion, educator or motivator. Although the effects of using an emotional expressions were minimal (which the authors explain by a ceiling effect), the emotional motivator performed better. Emotional speech is in general less intelligible and this was less of a problem for the motivator agent than for the educator role.

Summarising experiences gathered with their Roboceptionist system (Kirby et al., 2005), Kirby et al (2010) emphasise the importance of robots' adherence to human social norms for human-robot interaction to proceed smoothly, as exemplified by the use of expressive moods and emotions as integral part of social interaction. The importance of modelling long-term aspects of, and interactions between, emotions, moods, and attitudes results from the frequency with which human emotional reactions are caused by social interactions, influenced by societal and cultural norms, or used to communicate desires to other people (generating expectations). Emotions provide key contributions to forming common ground (Klein et al., 2004) and enabling effective communication; e.g., expression of the specific state of sadness may invite the specific desire to be comforted (cf. Feltovich et al.'s (2007) discussion of regulation devices in joint activity). Kirby et al. stress the importance of complete generative models of the interplay of different affective phenomena and the influence of interactions with people on them to be able to sustain long-term interactions. They propose a minimal generic design based on (discrete) *emotions*, *moods*, and *attitudes* and they mutual influences. Here, a basic distinction is made between *affect* as generalised state characterised by *valence* (positive or negative, or of approaching/engaging vs. distancing/disengaging (McNaughton, 2004)) from *emotion* as immediate response to some event as being of major significance. *Moods* are defined as more "diffuse" affective states of low variance (at least over a single day) and lower intensity than emotions and lacking in specific single antecedents. Finally, *attitudes*, acquired amalgamations of emotions experienced with a particular reference (person or object) over time, predispose to certain behaviours and imply evaluations on an affective scale (cf. the discussion of learning capabilities, below). First empirical tests of this model indicate the intelligibility of the robot's expressions, and how slight changes in them influence the evolution of interactions in terms of induced expectations and related coping strategies (avoidance of a negative robot, feeling of more common ground with a positive robot). Also, different levels of familiarity with the robot correlated with different reactions to robot's current mood. The proposed model was thus found capable of supporting rich and consistent long-term interactions with different human peers.

Guidelines

- **A (companion) robot/agent should express elements of its internal state (affective or otherwise). To find out which internal states are important and how they are expressed in a given situation, one can turn to studies of human-human expressions in similar situations, as expressivity is task/role/situation dependent.**
- **Principled generative models can provide important contributions by ensuring coherence while sustaining variability in long-term interactions.**

Situational Awareness

Most of the skills mentioned so far were concerned with the expressive elements of the robot or agent. For a conversation to work properly, one also has to be able to perceive and understand what the partner is saying, doing and expressing in general. For effective communication, the agent or robot needs to build up a model of the user's state for effective communication and of the situation. The motivation given by Castellano et al. (2008) is that companion robots should know when it is appropriate for the robot to engage in interaction with the user. For instance, when it is appropriate to interrupt the activity the user is engaged in. Situational awareness also includes recognition of affective and cognitive states (Castellano et al. 2008; Fong et al. 2003; Kidd 2008). The affect sensitivity should be robust and based on multimodal cues; it should not just include the prototypical emotions (Tomkins 1962, 1963; Ekman 1973, 1992) but also other, complex, social ones (as described by social constructivist or componential models of emotion, cf. e.g. Scherer 2000). It should also capture the dynamics of emotions (see also Kirby et al., 2010) in an interaction- and peer-dependent manner (see also previous paragraph). At the same time, the significant limitations of the current state-of-the-art in recognition "in the wild" (in particular of affect, but also of the very interaction peers) need to be considered.

Guidelines

- **Robots/agents should be designed so as to take into account the context of use (context-awareness, situation-awareness, user-awareness, affect-sensitiveness).**
- **If adequate, robust interpretation capabilities (for e.g. identification of the interacting peer, recognition of affective tone) are not available, then the actions of the robot/agent should be carefully designed to either match generic situations or carefully restricted contexts of use where the context can be validated.**

Summary

The premise that robots and agents should be human-like seems to be the dominant opinion of the field. However, as Kobyashi et al. (2008)'s successful example shows, this should not be mistaken for an immutable principle. When human-likeness is the goal, analysing human-human interactions in the selected domain and trying to copy the relevant behaviours on the system is a good way to proceed (see Bickmore's example for instance), but care must be taken to identify the full scope of the related required capabilities and to assess the viability of their technical realisation.

3.3 Social Skills

The dialogue and interaction skills presented above are needed to establish the desired interpersonal relation with the user. This relation, in turn, should be optimal to get closer to the overall goal of the agent: persuade the user to stop unhealthy habits, or advise on a healthier lifestyle, for instance. We list here some of the more important relational skills for companion or counselling agents as found in the literature.

Relational Communication Skills

Robots or agents that function as companions or health-advisors need to be able to communicate in ways that are particular to their role. Besides requirements that follow from the task (advise, persuade, counsel), many of the skills in the interaction can be modelled to relate to controlling interpersonal variables. These are important, precisely because of the particular task which affects people's self-image. Before we present some of these variables in the next paragraph, we point out some of the communication skills and interaction strategies that can support social interaction.

Bickmore et al (2009) argue that voluntary-use interfaces for long-term use require special consideration regarding user engagement to sustain possibly thousands of interactions. Engagement for a longitudinal application could be provided by having the agent tell personal stories that are entertaining and self-disclosing. Focussing on health-care agents, Bickmore (2002) points out that empathy and trust in particular are important relations between therapist and patient for the therapy to be effective. A similar focus on these interpersonal variables may be required for a coach-type agent or robot. He proposes, amongst others, the following set of strategies that should be varied over time for building the right rapport.

- Increasing breadth and depth of topics (mostly during social dialogue)
- Increasing amount of small talk
- Increasing amount of information the agent knows about the user (telling the agent things is a type of "investment" made in the relationship)
- Increasing use of empathy, agreement with the user, and humour
- References to past interactions and mutual knowledge should increase over time
- References to future interactions, inclusive pronouns, expressing happiness to see the user, nonverbal immediacy behaviours, and the use of greeting routines, forms of address.

Some of these strategies can be used to implement personalisation and adaption to the user, which are important elements in longer relationships (see below).

According to Bickmore, the most important strategies to use in this type of interaction are:

- Meta-relational communication: "being very clear up front about the roles and expectations of each of the parties in the relationship, and checking in from time to time to see how everything is going and making adjustments as needed"
- Appropriate use of politeness and facework
- Appropriate use of empathy: this "can go a long way towards making them feel understood and alleviating negative emotional states such as frustration" (p.198, with reference to (Klein et al, 2002))

Blanson-Henkemans et al. (2009) emphasise the need for feedback for a persuasive computer assistant. In their experiment they had the iCat follow principles of motivational interviewing, focusing on social functioning: discussing problems and giving feedback in the form of advice and direction. They implement strategies such as the following:

- Express empathy: "You did not achieve your goal, perhaps you were busy."
- Cheering and Complementing: "Well done!"
- Support self-efficacy and optimism: "You did not achieve your goal. Don't worry. It will go step by step."

The strategies chosen clearly depend on the framework for intervention that is chosen.

Interpersonal variables

Rapport and empathy were two interpersonal variables mentioned in the previous paragraph that the relational skills are intended to ensure. One could say that a general requirement for social companions is the ability to engage in the actions and strategies of the previous paragraph to set, maintain, or increase the value of these and several other similar variables mentioned in the literature mentions. Kidd (2008), for instance, stresses engagement, motivation to use the system (providing encouragement), and trust. Looije et al. (2006) state that a personal assistant promoting health should be able to:

- have users like it (so that they will engage with it): involving relational strategies such as giving advice in a positive manner; giving emotional support (empathy and compassion);
- have users trust it.

In Bickmore's work on FitTrack (Bickmore 2002), the main dimension of the relationship is the "working alliance" which is based "on the trust and belief that the therapist and patient have in each other as team-members in achieving a desired outcome". Crucial in forming and maintaining the working alliance is the patient's perception of empathy.

Guideline

- **Use the appropriate conversational strategies to establish the right settings for important relational variables.**

Personalisation and Adaptation / Memory and Learning

Strategies to maintain a relationship and produce the desired effect – be it companionship or the promotion of a healthier life style – may need to be varied depending on the person (Green et al. 2004, for instance) and on the particular moment. Several authors point out that the ability to adapt to a specific user over time – personalisation – is an important prerequisite. It will – as was pointed out above – require of an agent that it can build up a model of the user (personality, desires, preferences, attitudes, beliefs, etc.) and update it dynamically (Bickmore (2002), Castellano et al. (2008), Kidd (2008)). The reverse side of this is that also agents themselves need to be consistent (though they can also change over time) and have some kind of personality that is consistent with their role (cf. the generative model in (Kirby et al., 2010), discussed above). Bickmore et al. (2009) investigated the use of autobiographical stories for agents as a form of self-disclosure.

Another aspect closely related to time awareness is the capability for intelligent social Companions to remember and forget, as has been claimed by some authors. A companion should be able to remember past experiences to personalise the interaction and maintain a long-term relationship. On the other hand, the companion should also be able to forget (situational forgetting). One motivation behind this involves the ability to protect the user's privacy by not disclosing sensitive data. In general it is claimed that memory is an important aspect of social life. Memory is important for consistent action and thereby defining a personality (Lim et al., 2009). Castellano et al. (2008) point out that a memory model is needed to retrieve memories that are relevant to the current situation. The memory should help in ranking the significance of events and provide a basis for learning. Believability and richness of the interactive experience are claimed to benefit from this, in accord with current theorising in personality research (Mischel, 2004).

Guideline

- **Work on memory systems – in particular forgetting – in long-term interactions is fairly new ground for the research community. Incorporating the adequate strategies for user modelling is an important requirement for continuous engagement and attachment.**

Summary

For companion, health advising, therapist and other such agents and robots, the interpersonal variables between system and user have to be "right"; just as they have to be in real-life human-human interaction. This requires special interactive skills.

4 Summary and Conclusions

The studies on human-robot and human-agent interactions have interesting insights to offer that may be useful for the design of a socially engaged robot. However, the literature does not offer concrete, ready-to-use solutions to apply in every case. Rather, it articulates concerns, stipulates requirements or wish lists, reports on detailed experiments that may not be transferable to other situations, offers some insights gained in designing, implementing and evaluating an artificial social system, and provides suggestions for the design process. The table below summarises the main types of findings from this overview.

Appearance	1. Physical robots are preferred over virtual agents 2. Agents are preferred over text 3. Appearance is not as important as function 4. Appearance should match task	The literature on appearance consists mainly of laboratory experiments in which different appearances are compared. Most relevant for SERA are findings 3 and 4.
Basic Communicative Skills	1. Robots and agents should to a large extent communicate multi-modally just as humans can. 2. Robots and agents can use other means to communicate.	There is ample literature on how conversational skills that humans employ could be modelled in embodied dialogue systems and what the effects are on the quality of conversation and on the social impact (see, e.g. the Intelligent Virtual Agent and AAMAS conference series). In general, users will read and expect similar messages from the verbal and nonverbal behaviours of robots and agents as they do from humans; and ascribe personalities and social identities.
High level Skills	1. Goals for social agents on a higher level can be defined as engaging in behaviours that promote the right setting of the proper interpersonal variables. 2. Recurrent interpersonal variables are trust and empathy. 3. The intervention method chosen determines the task domain and the strategies that are important; as well as the interpersonal variables that matter.	The literature offers some good examples of how a specific intervention not only leads to a choice of dialogue strategies, dialogues acts and important expressive behaviours (e.g. by analysing the human-human interventions) but also the effects, variables and higher level skills that are important to model.

One final aspect to consider with respect to the literature and these guidelines is how they relate to the goals of the SERA project. As our aim is to make progress towards social robots, also by providing guidelines for improving sociability of this and the next generation of robots, the suggestions made in the literature on how to keep users engaged for longer time are central. In particular for the purposes of the in situ studies of human-robot interaction in WP2, they may facilitate the capturing of sufficient quantities of relevant data.

As mentioned in the opening, this deliverable complements D1.2a from the theory workpackge WP1. D1.2a mainly reviewed social psychological theories on human-human interaction and concluded with general guidelines for human-agent/robot interaction that were extrapolated to this case. In contrast, the current deliverable (D 3.1) starts from the description and evaluation of agents and robots that have been built and evaluated. The focus thus is less on *what* should be achieved but *how* this can be achieved, i.e., sociability through appearance and specific kinds of communicative and other behaviours. There are several points where the two deliverables meet, directly or indirectly. It is clear that many of the systems that have been created consider similar issues of social psychology and provide ideas on how to realise communication and create social bonds for which Theory-of-Mind like skills and the need-to-belong concepts of the theory framework are a prerequisite. Deliverable 3.2 on the Reference Architecture will take a step further by looking at the implementation details of the robots and agents that have been created.

With respect to the construction and evolution of the set-up with the Nabaztag that will be used in the data collection iterations, the present analysis of existing systems can make three kinds of contributions: The first are insights into issues that need to be considered in the design of the set-up. The second is the presentation of a number of methods for the design. The third are suggestions for appearance and behaviours. Given the limited state-of-the-art in human-robot/agent interaction, the fact that the research is still very much ongoing, and because of the overall complexity of the task, it is of little surprise that the insights gained from the literature are often not of an immediate prescriptive kind. Results may not generalize, and substantial empirical evidence is still missing for a number of ideas. All the more important are also the lessons taught by explicitly reported negative results and by the stumbling blocks encountered during the evolution of systems. The aim of the empirical work in SERA set-ups is not to build *the* ultimate social robot companion, but to validate existing knowledge and to gather empirical evidence that is currently missing. Importantly, this also includes making explicit implementation-time decisions that are all too often glossed over in scientific publications[3].

[3] See the benefit of “Clarity and completeness” of computational cognitive modelling discussed in (Cooper et al., 1996) and (Fum et al., 2007))

References

E. André, L. Dybkjær, W. Minker, and P. Heisterkamp (eds.), *Affective Dialogue Systems, Tutorial and Research Workshop (ADS 2004)*, Kloster Irsee, Germany, June 14-16, 2004, Proceedings, Lecture Notes in Computer Science, Volume 3068, Springer Berlin Heidelbery, 2004.

E. André, M. Rehm, W. Minker, and D. Bühler, Endowing Spoken Language Dialogue Systems with Emotional Intelligence, in (André et al. 2004), pp. 178-187, 2004..

C. Bartneck, Interacting with an Embodied Emotional Character, in J. Forlizzi, and B. Hanington (eds.), *Proceedings of the 2003 international conference on Designing pleasurable products and interfaces*, ACM Press, New York, NY, pp. 55-60, 2003.

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

T. W. Bickmore, L. Caruso, K. Clough-Gorr, and T. Heeren, \`It's just like you talk to a friend' -— Relational Agents for Older Adults, *Interacting with Computers*, 17(6):711- 735, 2005.

T. W. Bickmore and R. W. Picard, Establishing and Maintaining Long-Term Human-Computer Relationships, *ACM Transactions on Computer Human Interaction*, 59(1):21-30, 2005.

T. W. Bickmore, D. Schulman, and L. Yin, Engagement vs. Deceit: Virtual Humans with Human Autobiographies, in (Ruttkay et al. 2009), pp. 6-19, 2009.

O. A. Blanson Henkemans, P. J. M. van der Boog, J. Lindenberg, C. A. P. G. van der Mast, M. A. Neerincx, and B. J. H. M. Zwetsloot-Schonk, An online lifestyle diary with a persuasive computer assistant providing feedback on self-management, *Technology and Health Care*, 17(3):253-267, 2009.

C. Breazeal, C. D. Kidd, A. L. Thomaz, G. Hoffman, and M. Berlin, Effects of nonverbal communication on efficiency and robustness in human-robot teamwork, in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2005*, Edmonton, Alberta, Canada, August 2005. IEEE Press, Washington, D.C., pp. 708-713, 2005.

A. Bruce, I. Nourbakhsh, and R. Simmons, The role of expressiveness and attention in human-robot interaction, in *Proc. IEEE International Conference on Robotics and Automation (ICRA '02)*, vol.4, IEEE Press, Washington, D.C., pp. 4138-4142, 2002.

D. Byrne, G. Clore, and G. Smeaton, The attraction hypothesis: do similar attitudes effect anything? *Journal of Personality and Social Psychology* 51:1167–1170, 1986.

J. Cassell and T. W. Bickmore, Negotiated Collusion: Modeling Social Language and its Relationship Effects in Intelligent Agents, *User Modeling and User-Adapted Interaction*, 13(1-2): 89-132, 2003.

J. Cassell, T. W. Bickmore, M. Billinghurst, L. Campbell, K. Chang, H. H. Vilhjálmsson, and H. Yan, Embodiment in Conversational Interfaces: Rea, in *Proc. CHI, 1999*, ACM Press, New York, NY, pp. 520-527, 1999.

G. Castellano, R. Aylett, K. Dautenhahn, A. Paiva, P. W. McOwan, and S. Ho, Long-term affect sensitive and socially interactive companions, in *Proceedings of the Fourth International Workshop on Human-Computer Conversation*, 2008.

G. Castellano, I. Leite, A. Pereira, C. Martinho, A. Paiva, and P. W. McOwan, It's All in the Game: Towards an Affect Sensitive and Context Aware Game Companion, in (Cohn et al., 2009) pp. 29-36, 2009.

J. Cohn, A. Nijholt, and M. Pantic (eds.): *Proceedings of the International Conference on Affective Computing and Intelligent Interaction (ACII 2009)*, September 10-12, 2009, Amsterdam, The Netherlands, Vols I & II, IEEE Press, Washington D.C., 2009.

G. Clarizio, I. Mazzotta, N. Novielli, and F. de Rosis, Social Attitude Towards A Conversational Character, in *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2006)*, IEEE Press, Washington, D.C., pp. 2-7, 2006.

R. Cooper, J. Fox, J. Farringdon, and T. Shallice, A systematic methodology for cognitive modelling. Artificial Intelligence, 85(1-2):3-44, 1996.

K. Dautenhahn, Design spaces and niche spaces of believable social robots, in *IEEE Int. Workshop on Robot and Human Interactive Communication*. Berlin, Germany, IEEE Press, Washington, D.C., pp.192-197, 2002.

K. Dautenhahn, S. N. Woods, C. Kaouri, M. L. Walters, K. L. Koay, and I. Werry, What is a robot companion—friend, assistant or butler? in *Proceedings of IEEE RSJ international conference on intelligent robot systems (IROS'05)*, IEEE Press, Washington, D.C., pp. 1488-493, 2005.

B. R. Duffy, Anthropomorphism and the social robot, *Robotics and Autonomous Systems*, 42(3-4):177-190, 2003.

P. Ekman, Darwin and cross-cultural studies of facial expression, in P. Ekman (ed.): *Darwin and Facial Expression – a century of research in review*, Academic Press, New York, pp. 1-83, 1973.

P. Ekman, Basic emotions, in T. Dalgleish and M. Power (eds.), *Handbook of Cognition and Emotion*, Wiley, Chichester/London/New York, pp.45-60, 1999.

P. J. Feltovich, J. M. Bradshaw, W. J. Clancey, and M. Johnson, Toward an Ontology of Regulation: Socially-Based Support for Coordination in Human and Machine Joint Activity, in G. M. P. O'Hare et al. (eds.) Engineering Societies in the Agents World VII, Lecture Notes in Computer Science, Volume 4457, Springer Berlin Heidelberg, pp. 175-192, 2006.

B. J. Fogg, Creating persuasive technologies: an eight-step design process, in S. Chatterjee and P. Dev (eds.): *Persuasive Technology, Fourth International Conference (PERSUASIVE 2009), Proceedings*, ACM Press, New York, NY, article 44, 2009.

B. J. Fogg, and C. Nass, Silicon sycophants: the effects of computers that flatter, *International Journal of Human-Computer Studies*, 46(5):551-561, 1997.

T. Fong, I. Nourbakhsh, and K. Dautenhahn, A survey of socially interactive robots, *Robotics and Autonomous Systems*, 42(3-4):143-166, 2003.

D. Fum, F. D. Missier, and A. Stocco, The cognitive modeling of human behavior: Why a model is (sometimes) better than 10,000 words. *Cognitive Systems Research*, 8(3), 135-142, 2007.

J. Goetz, S. Kiesler, and A. Powers, Matching robot appearance and behavior to tasks to improve human-robot cooperation, in *Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2003)*, IEEE Press, Washington, D.C., pp. 55-60, 2003.

K. Göttlicher, S. Stein, and D. Reichardt, Effects of Emotional Agents on Human Players in the Public Goods Game, in (Cohn et al. 2009), pp.656-661, 2009,

N. Green, W. Lawton, and B. Davis, An assistive conversation skills training system for caregivers of persons with Alzheimer's Disease, in T. Bickmore (ed.) *Dialogue Systems for Health Communication, Papers from the 2004 AAAI Fall Symposium*, Technical Report FS-04-04, AAAI Press, Menlo Park, CA, pp. 36-43, 2004.

J. Grolleman, B. van Dijk, A. Nijholt, and A. van Emst, Break the Habit! Designing an e-Therapy Intervention Using a Virtual Coach in Aid of Smoking Cessation, in IJseelsteijn W. et al. (eds.) *Persuasive Technology*, Lecture Notes in Computer Science, Vol. 3962, Springer Berlin Heidelberg, pp. 133-141, 2006.

V. Groom, C. Nass, T. Chen, A. Nielsen, J. K. Scarborough, and E. Robles, Evaluating the effects of embodied realism in embodied agents. *International Journal of Human-Computer Studies* , 67(10): 842-849, 2009.

F. Hegel, M. Lohse, and B. Wrede, "Effects of visual appearance on the attribution of applications in social robotics," in *Proc. of the 18th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2009)*, IEEE Press, Washington, D.C., pp. 64-71, 2009.

R. C. Hubal, C. I. Guinn, D. C. Sparrow, E. J. Studer, R. D. Day, and W. A. Visscher, A Synthetic Character Application for Informed Consent, in T. Bickmore (ed.) *Dialogue Systems for Health Communication, Papers from the 2004 AAAI Fall Symposium*, Technical Report FS-04-04, AAAI Press, Menlo Park, CA, pp. 58-63, 2004.

E. Hudlicka, S. Payr, R. Ventura, C. Becker-Asano, K. Fischer, I. Leite, A. Paiva, and C. von Scheve, Social Interaction with Robots and Agents: Where Do We Stand, Where Do We Go? In: (Cohn et al. 2009), pp. 698-703, 2009.

D. Johnson, J. Gardner, and J. Wiles, Experience as a moderator of the media equation: the impact of flattery and praise. *International Journal of Human-Computer Studies* 61(3):237-258, 2004.

D. Johnson, and J. Gardner, The media equation and team formation: Further evidence for experience as a moderator. *International Journal of Human-Computer Studies*, 65(2):111-124, 2007.

T. Kanda, T. Myiashita, T. Osada, Y. Haikawa, and H. Ishiguro, Analysis of humanoid appearances in human-robot interaction, *IEEE Transaction on Robotics* 24(3):725-735, 2008.

J. M. Kessens, M. A. Neerincx, R. Looije, M. Kroes, and G. Bloothooft, Facial and Vocal Emotion Expression of a Personal Computer Assistant to Engage, Educate and Motivate Children, in (Cohn et al. 2009), pp. 621-627, 2009.

Z. Khan, Attitudes towards intelligent service robots, Technical report. IPLab, NADA, Royal Institute of Technology. TRITA-NA-E98421, IPLab-154, 1998.

C. D. Kidd, Designing for Long-Term Human-Robot Interaction and Application to Weight Loss, PhD Thesis, Massachusetts Institute of Technology, School of Architecture and Planning, Program in Media Art and Sciences, 2008.

R. Kirby, F. Broz, J. Forlizzi, M. P. Michalowski, A. Mundell, S. Rosenthal, B. P. Sellner, R. Simmons, K. Snipes, A. Schultz, and J. Wang, Designing Robots for Long-Term Social Interaction, in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005)*, IEEE Press, Washington, D.C., pp. 2199-2204, 2005.

R. Kirby, J. Forlizzi, and R. Simmons, Affective social robots. *Robotics and Autonomous Systems*, 58(3):322-332, 2010.

G. Klein, P. J. Feltovich, J. M. Bradshaw, and D. D. Woods, Common ground and coordination in joint activity, in W. B. Rouse and K. R. Boff (eds.), Organizational Simulation, John Wiley & Sons, New York, NY, pp. 139-184, 2004.

K. L. Koay., D. S. Syrdal, M. L. Walters, and K. Dautenhahn, Living with robots: Investigating the Habituation Effect in Participant's preferences during a longitudinal human-robot interaction study. *Proceedings of the 16th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN 2007)*, IEEE Press, Washington, D.C., pp. 564-569, 2007.

J. Klein, Y. Moon, and R. Picard, This Computer Responds to User Frustration: Theory, Design, Results, and Implications. *Interacting with Computers*, 14:119-140, 2002.

K. L. Koay, D. S. Syrdal, M. L. Walters, and K. Dautenhahn, Five weeks in the robot house – Exploratory human-robot interaction trials in a domestic setting, *Second international conference on advances in computer-human interaction*, IEEE Press, Washington, D.C., pp. 219-226, 2009.

K. Kobayashi, K. Funakoshi, S. Yamada, M. Nakano, Y. Kitamura, and H. Tsujino, Smoothing human-robot speech interaction with blinking-light expressions, in *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008)*, IEEE Press, Washington, D.C., pp. 47-52, 2008.

H. Kozima, C. Nakagawa, and H. Yano, Attention coupling as a prerequisite for social interaction, in *Proc. of the 12th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2003)*, IEEE Press, Washington, D.C., pp. 109-114, 2003.

I. Leite, A. Pereira, C. Martinho, and A. Paiva, Are emotional robots more fun to play with? in *Proc. of the 17th IEEE International Sympoium on Robot and Human Interactive Communication (RO-MAN 2008)*, IEEE Press, Washington, D.C., pp. 77-82, 2008.

S. Li, B. Wrede, and G. Sagerer, A dialog system for comparative user studies on robot verbal behavior, in *Proc. of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2006)*, IEEE Press, Washington, D.C., pp. 129-134, 2006.

M. Y. Lim, R. Aylett, W. C. Ho, S. Enz, and P. Vargas, A Socially-Aware Memory for Companion Agents, in (Ruttkay et al. 2009), pp. 20-26, 2009.

R. Looije, F. Cnossen, and M. A. Neerincx, Incorporating guidelines for health assistance into a socially intelligent robot, in *Proc. of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2006)*, IEEE Press, Washington, D.C., pp. 515-520, 2006.

R. Looije, M. A. Neerincx, and F. Cnossen, Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors, *International Journal of Human-Computer Studies*, vol. in Press, Corrected Proof, 2009.

K. F. MacDorman, Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: an exploration of the uncanny valley, in *Proceedings of the ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science*, pp. 26–29, 2006.

N. McNaughton, The conceptual nervous system of J. A. Gray: Anxiety and neuroticism. *Neuroscience & Biobehavioral Reviews*, 28(3):227-228, 2004.

W. Mischel, Toward an integrative science of the person. *Annual Reviews in Psychology*, 55:1-22, 2004.

T. Nomura, T. Tasaki, T. Kanda, M. Shiomi, H. Ishiguro, and N. Hagita, Questionnaire-based social research on opinions of Japanese visitors for communication robots at an exhibition. *AI & Society* 21:167-183, 2007

T. Nomura., T. Suzuki, T. Kanda, J. Han, N. Shin, J. Burke, and K. Kato, What people assume about robots: cross-cultural analysis between Japan, Korea and the USA. *Human-robot interaction.* N. Sarkar. Vienna, Austria, Itech education and publishing, 275-288, 2007.

A. Oyedele, S. Hong, and M. S. Minor, Contextual factors in the appearance of consumer robots: exploratory assessment of perceived anxiety toward human-like consumer robots, *CyberPsychology and Behavior 10(5)*, 624-632, 2007.

M. E. Pollack, L. Brown, D. Colbry, C. Orosz, B. Peintner, S. Ramakrishnan, S. Engberg, J. T. Matthews, J. Dunbar-Jacob, C. E. McCarthy, S. Thrun, M. Montemerlo, J. Pineau, and N. Roy, Pearl: A mobile robotic assistant for the elderly, in K. Haigh (ed.) Automation as Caregiver: The Role of Intelligent Technology in Elder Care, Papers from the AAAI Workshop, Technical Report WS-02-02, AAAI Press, Menlo Park, CA, pp.85-91, 2002.

B. Reeves and C. Nass, *The Media Equation: how people treat computers, televisions and new media like real people and places*. Cambridge, Cambridge University Press, 1996.

B. Robins, K. Dautenhahn, R. te Boekhorst, and A. Billard, Robots as assistive technology – Does appearance matter? *Proceedings of the 13th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2004)*, IEEE Press, Washington, D.C., 277-282, 2004.

J. Robison, S. McQuiggan, and J. Lester, Evaluating the Consequences of Affective Feedback in Intelligent Tutoring Systems, in (Cohn et al., 2009), pp. 37-42, 2009.

F. de Rosis, N. Novielli, V. Carofiglio, A. Cavalluzzi, and B. D. Carolis, User modeling and adaptation in health promotion dialogs with an animated character, *Journal of Biomedical Informatics* 39(5):514- 531, 2006.

M. Rumpler, EXSTASIS – An Extended Status Model for Social Interactions, in (Ruttkay et al., 2009), pp. 49-55, 2009.

Z. Ruttkay, M. Kipp, A. Nijholt, and H. H. Vilhjálmsson (eds.) *Intelligent Virtual Agents, 9th International Conference, IVA 2009*, Amsterdam, The Netherlands, September 14-16, 2009, Proceedings, Lecture Notes in Computer Science, Vol. 5773, Springer, Berlin Heidelberg, 2009.

Scherer K.R., Psychological models of emotion, in J. Borod (ed.): *The neuropsychology of emotion*, Oxford University Press, Oxford New York, pp.137-166, 2000.

M. Scopelliti, M. V. Giuliani, A. M. D'Amico, and F. Fornara, If I had a robot at home… Peoples' representation of domestic robots, in: S. Keates, et al. (eds.) *Designing a more inclusive world*, Springer-Verlag, London, pp. 257-266, 2004.

P. J. Sengers, Designing Comprehensible Agents, in *Proceedings of the Sixteenth International Conference on Artificial Intelligence (IJCAI-99)*, Morgan Kaufmann, Los Altos/Palo Alto/San Francisco, pp.1227-1232, 1999.

T. Shibata, K. Wada, and K. Tanie, Tabulation and analysis of questionnaire results of subjective evaluation of seal robot at Science Museum in London, in *Proceedings of the 11th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2002)*, IEEE Press, Washington, D.C., pp. 23-28, 2002.

C. L. Sidner, C. D. Kidd, C. Lee, and N. Lesh, Where to look: a study of human-robot engagement, in *Proceedings of the 9th international conference on Intelligent user interfaces (IUI'04)*, ACM Press, New York, NY, USA, pp. 78-84, 2004.

C. L. Sidner, C. Lee, C. D. Kidd, N. Lesh, and C. Rich, Explorations in engagement for humans and robots, *Artificial Intelligence*, 166:140-164, 2005.

A. Tapus, C. Tapus, and M. J. Mataric, The role of physical embodiment of a therapist robot for individuals with cognitive impairments, in *Proc. of the 18th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2009)*, IEEE Press, Washington, D.C., pp. 103-107, 2009.

S. S. Tomkins, Affect, Imagery, Consciousness, Vol.1: The Positive Affects, Springer, New York, 1962.

S. S. Tomkins, Affect, Imagery, Consciousness, Vol.2: The Negative Affects, Springer, New York, 1963.

J. Wainer, D. J. Feil-Seifer, D. A. Shell, and M. J. Mataric, The role of physical embodiment in human-robot interaction, in *Proc. of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2006)*, IEEE Press, Washington, D.C., pp. 117-122, 2006.

M. L. Walters, D. S. Syrdal, K. Dautenhahn, R. te Boekhorst, K. L. Koay, Avoiding the uncanny valley: robot appearance, personality and consistency of behaviour in an attention-seeking home scenario for a robot companion. *Autonomous Robots* 24(2):159-178, 2008.