The Case of K.I.T.T. and Data – from Science Fiction to Reality?

A Social Psychology Perspective on Artificial Companions

Astrid M. Rosenthal-von der Pütten (University of Duisburg-Essen, a.rosenthalvdpuetten@uni-due.de)

Nicole C. Krämer (University of Duisburg-Essen, nicole.kraemer@uni-due.de)

Abstract

The present paper aims to provide a state-of-the-art overview of research on artificial companions from a social psychology perspective. More specifically, it follows two objectives: First, it outlines a theoretical framework of sociability in which concepts and theories from social psychology are organized in a three-level model. The concepts and theories introduced are discussed with regard to their applicability to artificial companions on the basis of two companion examples from Science Fiction (K.I.T.T. and Data). In a résumé, the paper summarizes which concepts and theories are mandatory, useful, or marginally useful for the development of artificial companions, and which concepts are limited in their explanatory power. Second, the paper provides an overview on current artificial companion research and outlines corresponding methodological challenges. Various subjective and objective measures are introduced. The need for a multi-method approach and long-term studies is discussed.

1 Introduction

K.I.T.T.: Michael, why do you need to socialize with so many women? Wouldn't one be sufficient?

Michael: K.I.T.T., you're beginning to sound like my mother, here. I mean, what's wrong with a little companionship?

K.I.T.T.: Eh?

Michael: You can understand that.

K.I.T.T.: No, Michael, I cannot. When you're one-of-a-kind, companionship does not compute.

With the development of companion systems, research on virtual agents and robots gains increasing attention in the media and is brought to the public's focus. Indeed, social science research and public discussion on current developments is necessary since the implications can be discussed controversially. Do people want to share their bed and board with an artificial companion? In Science Fiction, companion technologies are part of the protagonists' daily lives. Michael Knight, for instance, has been teamed up with the robotic car K.I.T.T. and Commander Data is a well-respected member of the crew of the USS Enterprise. Our expectations on companion systems are greatly influenced by literature and movies starring full computerized environments like artificially intelligent houses or different kinds of mobile robots such as K.I.T.T. or Data. In the course of this paper these Sci-Fi companions will be used to exemplify a) the roles these systems take on, b) how they live and work together with humans, and c) problems this shared life entails. On that account we will shortly recap the design and features of the examples K.I.T.T. and Data.

In the TV series Knight Rider Michael Knight is teamed up with a supercomputer integrated in a Trans-Am sports car, the *Knight Industries Two Thousand* (K.I.T.T.). The Knight 2000 microprocessor as the core piece of K.I.T.T. includes the self-aware cybernetic logic module. Besides auto cruise, audio/video entertainment and surveillance capabilities, it features a computer voice with which K.I.T.T. is able to communicate via natural language. K.I.T.T. can collaborate, but also decide and act autonomously. His artificial intelligence is so advanced that he developed a kind of personality which can be characterized as benevolent and compassionate, but also sensitive and easily offended. In the course of the series, K.I.T.T. gradually forms relationships with Michael Knight and the other crew members. K.I.T.T. is programmed to protect human life, and thus he does not utilize lethal force. He uses a medical scanner to monitor vital signs of individuals and is able to identify whether people are injured, poisoned, undergoing stress or other emotional states (see http://knightrideronline.com and http://en.wikipedia.org/wiki/K.I.T.T.).

Lieutenant Commander Data - a fully functional android robot with a positronic brain - is the second officer of the starship USS Enterprise on the TV series Star Trek: The Next Generation. Data can be dis- and reassembled, does not need any life support to function (also under water, in different atmospheres or even in vacuum) and is immune to biological diseases. However, he can be affected by computer viruses, chip malfunctions and he can simply be switched off using a switch on his back. Data can be described as an emotionally handicapped robotic superhuman: On the one hand he looks stunningly human, is physically the strongest member of the crew, processes and calculates information as rapidly as a supercomputer. On the other hand, he cannot feel, is inured to sensory tactile feelings such as pain or pleasure and is unable to grasp basic emotions, imagination, and humour. Therefore, Data has on-going difficulties with understanding various aspects of human behaviour, but shows an aspiration to find his own humanity. Although Data is of mechanical nature, he is treated as an equal

member of the crew of the Enterprise (see also www.startrek.com; http://en.wikipedia.org/wiki/Commander_Data).

Despite the advanced robots in Science Fiction, the research realm of artificial companions is still in its infancy. Researchers across, but also within, disciplines do not necessarily agree on what exactly renders a technology an artificial companion (cf. Böhle & Bopp, this issue). Moreover, it is still hard to find meaningful fields of application for companion technologies which will be accepted by common users. Compared to the Sci-Fi examples of K.I.T.T. or Data, current companion technology is in its fledgling stages and far behind users' media-induced expectations on the abilities of companion technology. Some application fields, however, are useful for users and are also adequate test-beds to address a variety of different research questions. One of them is the health care sector where companions, e.g., serve as supervisor for physical activity for elderly people or post-stroke patients (von der Pütten et al. 2011b; Matarić et al. 2007), assist elderly or disabled people with everyday tasks at home (Kheng Lee Koay et al. 2009) or at work (Hüttenrauch et al. 2004), or support children with cognitive and physical disabilities (Robins et al. 2012). Other application fields also focus on target groups with special needs like elderly people who struggle with technology and could benefit from a more natural interaction with an embodied agent (Yaghoubzadeh 2011).

While the two exemplary Sci-Fi companions are perfectly designed systems users are happy to deal with, in reality researchers and developers face the frequently occurring phenomenon that people are initially interested in interacting with an artificial entity; but are, however, quickly bored or annoyed with it, refuse to use it again and even show aggression towards the system (de Angeli et

al. 2006; Walker et al. 2002). Nevertheless, embodied agents and other artificial entities were demonstrated to have positive emotional, cognitive and motivational effects. Diverse studies showed embodied agents to increase students' motivation to learn with tutoring programs (e.g., Krämer 2010; Lester et al. 2000; Eimler et al. 2010) and to improve students' learning performance (e.g., Baylor & Kim 2008; Eimler et al. 2010). Moreover, Krämer et al. (2003) demonstrated that participants were more forgiving and less negatively affected when a system failure was presented by an embodied TV-VCR agent compared to a text-based interface. These examples show the great potential of companion technologies such as virtual agents or robots to be beneficial in diverse tasks and for various target audiences. Thus, the central challenge is to further refine the sociability of artefacts that is considered to facilitate human-robot/agent interaction (HRI/HAI; Krämer et al. 2011).

Although it is difficult to draw conclusions regarding users' acceptance of future scenarios, we are able to address the question: What exactly makes a companion social? All systems presented allow examining aspects of sociability separately. In the first part of this paper, we will thus introduce a theoretical framework discussing several levels of sociability (see also Krämer et al. 2011). Based on the companion examples from Sci-Fi and state-of-the-art research we will critically reflect whether human-companion interaction has to build upon basic principles of human-human interaction or whether alternative approaches have to be considered.

A second major challenge in the research realm of artificial companions is to choose and use adequate methods to study human-companion relationships. Therefore, we will discuss the necessity for methodological interdisciplinarity, multi-method ap-

proaches and long-term (field) studies. Conducting field studies is difficult, because companion technologies are often not market-ready, the technical components are expensive or the system is error-prone and needs constant supervision. Moreover, analysing field data, especially from long-term studies, is highly time consuming and costly. Thus another major challenge for this research domain is to choose and use adequate methods to study human-companion relationships. In the second part of this paper, we will therefore provide an overview on methods used for artificial companion research and discuss their advantages, drawbacks and their feasibility on the basis of state-of-the-art research examples.

In sum, this paper will give an overview on existing research on companions in HCI and HRI, discuss the applicability of the underlying theoretical assumptions on the sociability of artefacts and provide an overview and discussion of methods used for artificial companion research.

2 Sociability of artificial entities – a three level model

Unanimously researchers agree that artificial entities which step in interaction with humans have to be sociable to facilitate human-artefact interaction (e.g., Breazeal 2002; Ishiguro 2006; Krämer et al. 2011). However, there is no consensus on what sociability means in terms of artificial artefacts and whether respective rules of sociability should be originated from human-human interaction. Addressing this debate from a social psychological point of view, Krämer, Eimler, von der Pütten and Payr (2011) introduced a theoretical framework discussing several levels of sociability in human-human interaction, their applicability for HRI and how useful they are as a starting point for a theoretical conceptualization of human-artefact interaction and relation-

ships. In the following we will a) briefly present the concepts within the defined three levels of sociability, and b) discuss the concepts on the basis of state-of-the-art research and two companion examples from Sci-Fi.

2.1 Three levels of sociability

Krämer et al. (2011) identify aspects of sociability which are organized and summarized in three different levels (see Table 1). In the present paper all three levels will be discussed on the basis of exemplary concepts within the respective level. For the discussion of all relevant concepts see Krämer et al. (2011).

On a micro-level, prerequisites for communication are addressed by demonstrating in which way Theory of Mind, perspective taking, and similar abilities enable social interaction. The meso-level contains concepts and theories from social psychology which describe the human need for relationships, what is needed to initially establish a relationship (e.g. reciprocity, attractiveness), and how it can be shaped and which factors affect their quality. On the macro-level, different roles are identified and discussed with regard to their helpfulness when trying to shape human-artefact interaction. Beyond addressing actual interaction and communication, the nature of the relationship and the role of the companion is discussed: should the relationship to the companion resemble an intimate long-term human-human relationship (e.g., family member, close friend), a non-intimate long-term human-human relationship (e.g., neighbour, mailman) or be rather based on human-pet relationships.

2.2 Micro-level: actual interaction & prerequisites for communication

According to Watzlawick, Beavin and Jackson (1967) people cannot not communicate. Any behaviour is a communicative act. Thus, in this paper, when speaking of interaction, in-

Table 1: Levels of Sociability

Levels of sociability	Corresponding theories
Micro-level: Actual interaction, Prerequisites for communication	• Common ground • Theory of Mind • Perspective Taking • Shared intentionality
Meso-level: Relationship building	• Need to belong • Prerequisites: mere exposure, attractiveness, reciprocity • Social exchange • Dimensions of human interaction will play a role (e.g. see dominance, intimacy)
Macro-level: Roles and persona	• Assignment of roles by designer versus user

teractive acts are interpreted as communicative acts. The focus of the micro-level of sociability lays on the prerequisites for communication. In this regard, the prerequisites common ground, Theory of Mind and perspective taking will be introduced and discussed. Although the three theories originated from different fields of research (communication science, ethology, cognitive science), they are to some extent overlapping concepts, all referring to the general ability of looking into someone's head. However, they are characterized by subtle differences and will therefore be discussed separately.

Common ground

K.I.T.T.: What does relax mean?
Michael: Um. It's kinda like when I put you in neutral.
K.I.T.T.: Oh. How very unproductive.

Common ground has been described as the joint basis for communication: ''Two people's common ground is, in effect, the sum of their mutual, common, or joint knowledge, beliefs, and suppositions'' (Clark 1992). The most obvious starting point in terms of communal common ground is human nature. As an example, Clark (1992) points out that if a sound is audible to someone, he will assume that it is audible to the other as well. Moreover, he explains that people take the same facts of biology for granted (e.g., everyone knows the bodily condition of being relaxed) and that everyone assumes certain social facts (people use language, live in groups, have names). It is obvious that artificial entities per default lack communal common ground unless the information is programmed (e.g. information on word meanings like "relax"). But providing the system with information on the biological nature of humans, their forms and rules of living together, does not imply that the system can make sense of this information.

Michael: K.I.T.T. I got a bone to pick with you.
K.I.T.T.: According to my data on human anatomy, you have 206 bones, give or take some questionable cartilage.

A human, even an individual from a different culture, would presumably be able to detect from the intonation of the sentence and by referencing to figurative language that Michael is not referring to an actual bone, but to an upcoming argument. If indeed in HHI the interlocutor fails to understand the contribution, humans still have verbal and nonverbal strategies to discover and repair situations. ''Contributors present signals to respondents, and then contributors and respondents work together to reach the mutual belief that the signals have been

understood well enough for current purposes'' (Clark 1992). Thus, feedback is a key concept also for human-artefact interaction, because it can compensate for a lack of knowledge. And further, learning can enhance system performance in a long-term view.

Perspective taking

[Michael talks to K.I.T.T. for the first time - very loudly and slowly]

K.I.T.T.: There's no reason for increased volume. I'm scanning your interrogatives quite satisfactorily. I am the voice of Knight Industry 2000's microprocessor, K.I.T.T. for easy reference.

The fact that the failure to take another's perspective into account can be the basis for misunderstandings and dispute, stresses the importance of perspective taking in human-human communication (see, e.g., Nickerson 1999; Rommeveit 1974). In this respect, a prerequisite for successful communication is that the message is tailored to the knowledge of the recipient (Krauss & Fussell 1991). Observing HRI/HAI, it is often found that users tailor their messages to the robot or agent and not the other way round - a phenomenon also known as computer-talk (Fischer 2006). Like Michael Knight, users speak more loudly, repeat themselves more slowly, or answer in a much simpler way than they would in human-human communication in order to compensate for technical shortcomings of the system. For instance Bell et al. (2003) demonstrated that speakers adapted their speech rate during interaction with an animated character. They spoke slower in response to a 'slow computer' and faster to a 'fast computer', respectively. This effect was mediated by overall performance of the system, e.g., when the computer seemed to have problems comprehending verbal input, participants speeded up less with the fast computer. Using discourse analysis, Shechtman et al. (2003) revealed a key difference in participants' behaviour in

HHI and HAI: When participants believed they were talking to a computer-mediated person instead of an artificial entity, they showed more of the kinds of behaviours associated with establishing the interpersonal nature of a relationship. However, the aim of companions is not to force the user to adapt to the system, but to allow natural interaction. Since perspective taking is a prerequisite for successful communication, also agents and robots should be able to tailor their messages to the user. This is often not realized in current systems. Moreover, when the human tries to compensate for the shortcomings of the system by adaptation, this is in most cases not successful as even basic concepts and -more importantly- contexts are not shared.

Theory of Mind

Lt. Jenna D'Sora: Kiss me. [Data obliges]

Lt. Jenna D'Sora: What were you just thinking?

Lt. Cmdr. Data: In that particular moment, I was reconfiguring the warp field parameters, analysing the collected works of Charles Dickens, calculating the maximum pressure I could safely apply to your lips, considering a new food supplement for Spot...

Lt. Jenna D'Sora: I'm glad I was in there somewhere.

The term ''Theory of Mind'' was coined by Premack and Woodruff (1978) as they referred to the ''ability –[...] to explain and predict the actions, both of oneself, and of other intelligent agents'' (Carruthers & Smith 1996). Theory of Mind (ToM) is the ability to see other entities as intentional agents, whose behaviours are influenced by states, beliefs, desires, etc. and the knowledge that other humans wish, feel, know, or believe something (Premack & Premack 1995; Premack & Woodruff 1978; Whiten 1991). Frith and Frith (2003) conclude that pragmatics of speech rely on mentalizing and that in many real-life cases the understanding of an utter-

ance cannot be based solely on the meanings of the individual words (semantics) or on the grammatical rules by which they are connected (syntax). Hence, humans go beyond the words we hear or read and hypothesize about the speaker's mental states. In the example presented above, Data fails to consider not only the actual words of the question Jenna asked, but to take the (to humans obviously romantic) situation and Jenna´s state and desires into account. If he had done so, he would have been able to infer that she did not want to hear about all actual computing processes going on in that particular moment, but some romantic answer solely referring to her and the kiss.

With regard to companion technologies, the obvious consequence of these considerations thus is to try to implement common ground, perspective taking, and Theory-of-Mind-like abilities, including the agent's ''awareness'' of its own abilities and the basic knowledge about the human interaction partner. However, as Frith and Frith (2003) aptly state, mere knowledge will not be enough to successfully mentalize: ''The bottom line of the idea of mentalizing is that we predict what other individuals will do in a given situation from their desires, their knowledge, and their beliefs, and not from the actual state of the world'' (Frith & Frith 2003: 6).

Nevertheless, Theory of Mind has been considered as a fruitful concept: "[…] a robot that can recognize the goals and desires of others will allow for systems that can more accurately react to the emotional, attentional, and cognitive states of the observer, can learn to anticipate the reactions of the observer, and can modify its own behaviour accordingly" (Scassellati 2002: 16). Recently, there are attempts to implement ToM-like abilities in agents (Peters 2006), robots (Breazeal et al. 2011), or multi-agent systems (Klatt et al. 2011). Krämer et al. (2011) presented a framework to

"demystify", i.e. to reduce the complexity of ToM abilities by distinguishing them on the basis of their properties (general vs. individual and static vs. dynamic properties) resulting in a matrix of ToM-abilities which makes it possible to analyse them and to design for them individually.

However, there is little known on how the implementation of ToM in artificial entities is perceived and evaluated by users. According to Waytz et al. (2010) the human brain is predestined to ascribe a mind to non-people under certain conditions such as social connection and similarity. Indeed, an fMRI experiment by Krach et al. (Krach et al. 2008) showed increased ToM-associated cortical activity in participants who completed a prisoner's dilemma task with game partners with increasing degrees of human-likeness (computer, a functional robot, an anthropomorphic robot, a human partner) regardless of the actual behaviour of the game partner which was completely random. Benninghoff et al. (2012) investigated whether implementing a Theory of Mind within a humanoid robot will lead to higher acceptance of the robot. They found that subjects acknowledged that a robot interacting with a human in a video showed Theory of Mind abilities, and rated the robot as more sympathetic and higher on social attractiveness. Yet it did not affect their evaluation of the robot's ability to fulfil a task satisfactorily.

Although it is assumed to bear great potential to facilitate human-artefact interaction, research and development is just at the outset of possibilities arising from the implementation of ToM-like abilities in artificial entities. Moreover, it can be debated whether applying the paradigm of human communication to companions is the right approach. While it might be regarded as advantageous that humans will not have to adapt in any way when they want to communicate with robots or virtual agents, it is obviously difficult

to implement crucial abilities for human-like communication. Alternatively, other communicative paradigms, like human-dog communication have been considered to be helpful models for human-robot/agent interaction (Dautenhahn 2004; Dautenhahn & Billard 1999) and have been implemented (Syrdal et al. 2010). Recent research suggests, however, that dogs also have several abilities that are not easily described by rules and are therefore not easy to implement. They are able to initiate communicative interactions, rely on visual human gestures, and recognize simple forms of visual (joint) attention (Miklósi 2009). It has been argued that dogs have been adapted to the human communication system by natural and breed selection (Tomasello 2008). Thus, the human-dog interaction model does not provide a more fruitful basis compared with human-human interaction, but a dog-shaped robot might induce lower expectations than a robot or agent with human-like appearance.

2.3 Meso-level: relationship building

The focus of the meso-level of sociability lays on the human need for and the establishment and maintenance of relationships. First, we will introduce humans' driving need to belong. Second, we will exemplify prerequisites for the establishment of relationships identified in social psychology research (e.g. attractiveness, reciprocity, propinquity) by outlining the importance of attractiveness. And third, we will address the topic of social equity which describes how relationships are negotiated and evaluated and its applicability with regard to artificial companions.

Need to Belong

K.I.T.T.: I hate to be the one to break this to you, but automobiles are not human. They have no lineage or personality.

Michael: I wonder why I keep forgetting that?

K.I.T.T.: You have probably begun to form a psychological attachment to me. That would be a logical human response.

K.I.T.T.'s statement that Michael's behaviour might be driven by the need of forming a psychological attachment indeed corresponds to human nature. Humans have been shown to possess a need to build relationships which has been termed the "need to belong" by Baumeister and Leary (1995) who suggest that "human beings are fundamentally and pervasively motivated by a need to belong, that is, by a strong desire to form and maintain enduring interpersonal attachments" (1995: 522). Thus, we seek the company of others in order to satisfy the need to belong. We build groups (e.g. families, cliques), help each other and join clubs just because the satisfaction of the need to affiliate makes us happy (see also Cacioppo & Patrick 2008; Ryan & Deci 2000). It has been claimed that humans are like "free monadic radicals" (Kappas 2005), eager to bond and affiliate with anything that is interactive and provides basic social cues such as, for example, speech (see Reeves & Nass 1996; Nass & Moon 2000). Indeed, a longitudinal study within the EU project SERA (Social Engagement with Robots and Agents) showed that some people established a kind of relationship with a robotic supervisor for physical activity placed in their house (SERA), including giving it a name, talking to it although it did not understand natural speech and stating to miss it after it was taken away from participants (von der Pütten et al. 2011b). Similar observations have been made for robotic pets (Fernaeus et al. 2010; Joana Dimas et al. 2010) and domestic devices like vacuum cleaners (Sung et al. 2010; Forlizzi 2007). However, throughout these studies not all participants showed attachment, and those who did showed different degrees of attachment. Thus, it is important to acknowledge the fact that in human-human interaction, humans will not just bond with any entity when given the choice, but that there

are factors that influence who is perceived to be attractive and whom we choose for the establishment of a relationship (see Aronson et al. 2010) which will be discussed in the following.

Attractiveness

Lt. Cmdr. Data: Darling, you remain as aesthetically pleasing as the first day we met. I believe I am the most fortunate sentient in this sector of the galaxy.

It can be assumed that humans will draw on similar criteria as they would in human-human encounters, when deciding whether they would like to interact again with a robot. In this regard, (physical) attractiveness plays an important role. Here, the finding ''what is beautiful is good'' (Dion et al. 1972), in the sense that attractive people are also rated positively in other aspects, can also be assumed to be true for agents and robots. It has been shown that the same principles for judging the attractiveness of humans hold for the judgment of attractiveness for virtual agents (Sobieraj 2012). Von der Pütten and Krämer (2012) identified different characteristics of robot appearances (e.g., mechanical/ humanoid/ android, but also toy-like and colours) which resulted in different ratings of the robots with regard to their likability. Thus, we know that artificial entities follow the same principles of physical attractiveness when they expose a humanlike appearance like Data. However, there is still little known on what exactly is perceived as beautiful when it comes to robots which are not android.

As an additional factor for relationship building, reciprocal liking might be taken into account. Since all humans like to be liked, we are attracted to others who behave as if they like us (Berscheid & Walster 1978; Kenny 1994; Kubitschek & Hallinan 1998). Liking can even compensate the absence of similarity (Gold et al. 1984). There are relatively easy ways to exploit reciprocal liking: that is the robot should give its user the impression that it likes him or her and appreciates his or her presence since this increases the likeability of the system. Depending on the setting, this may well be realized with the help of ingratiation (i.e., by praising the user). But it is important not to rely too much on seemingly simple, straightforward rules that are derived, because positive feedback and friendly behaviour is not always perceived positively, since, e.g., persons with a negative self-concept tend not to respond to the friendly behaviours of others and will provoke negative reactions affirming their negative self-concept instead (Swann et al. 1992).

Theories of social exchange and equity

Lt. Jenna D'Sora: This is all part of a program?

Lt. Cmdr. Data: Yes. One which I have just created for romantic relationships.

Lt. Jenna D'Sora: So I'm, erm... I'm just a small variable in one of your new computational environments?

Lt. Cmdr. Data: You are much more than that, Jenna. I have written a subroutine specifically for you - a program within the program. I have devoted a considerable share of my internal resources to its development.

Lt. Jenna D'Sora: Data... that's the nicest thing anybody's ever said to me.

The social exchange theory (Homans 1961; Thibaut & Kelley 1959) assumes that relationships are comparable to a marketplace where costs and benefits are exchanged according to economic principles. It can be summarized as ''the idea that people's feelings about a relationship depend on their perception of the rewards and costs of the relation, in the kind of relationship they deserve, and their chances of having a better relationship with someone else'' (Aronson et al. 2010). Hence, a person's level of satisfaction in a relationship is determined by the comparison level (Kelly & Thibaut

1978). The comparison level refers a) to the expected outcome of rewards and punishments the person is likely to receive in a relationship compared to previous experiences, b) the benefits and costs of alternative relationships, and c) the perception of how likely one could find an alternative partner to replace the old relationship. In the example of Data and Jenna, Jenna receives full attention by Data who wrote subroutines particularly for her. However, compared to previous and potential alternative relationships, she might experience less intimacy and emotional affection from her boyfriend. The question arises whether humans tend to compare a relationship with an artificial entity with the cost and rewards invested in "real" human-human relationships, or if other rules are applied. Also, it has to be considered to what kind of relationships the relationship with a robot/agent is compared: An adult, a child or, say, a pet. Considering the latter, many people have intense relationships with their dogs or cats although these animals can neither speak nor do they have any concept of human communication. Thus, the emotional rewards people gain seem to outweigh the costs they invest (e.g., food, medical care, time). Unlike these animals, robots are no living creatures, they are not warm and do not (at the moment) make the impression of acting autonomously. However, the data from the SERA project show that people were influenced by a robot's presence, at least; they felt that there was "something" (von der Pütten et al. 2011b). Additionally, Kahn et al. (2012) showed that children interacting with the robot Robovie believed that Robovie should not be harmed psychologically (although it could be bought and sold). Thus, if future research shows that humans build bonds that will lead them to feel sorry for the ending of the relationship with a robot/ agent, of course ethical questions will have to be discussed.

2.4 Macro-level: persona & roles

K.I.T.T.: I am still learning about the complexities of friendship, but I would be honoured to count you as mine.

Like many areas presented previously, there are also very few studies addressing possible personas and roles for companions. Robots in Sci-Fi are predominantly depicted as valuable and most of the time equally treated team members with some sort of personality. K.I.T.T. and Data both fulfil certain roles based on human role models (team/crew member, friend, boss). Unlike in Science Fiction, interviews on robots in real life, however, show that people - although generally in favour of a robot companion - saw its potential role as being an assistant, machine, or servant and only a few expressed the wish that the robot companion might be a friend (Dautenhahn et al. 2005). In sum, less intimate social roles or personalities were discussed, such as a butler or maid personality, a health adviser or a manager (for a specific part of the user's life). All of these social roles were associated with different capabilities of the system and expectations on behalf of the user. However, empirical research showed that the human user defines the way she/he perceives the robot/agent, the way she/he communicates with the robot/agent, and which role she/he assigns to the artificial entity (e.g., von der Pütten et al. 2011b; see also the results of the media equation, Reeves & Nass 1996). Thus, the perception of the robot/agent and its assigned role can be very different from the perception and role intended by the developer of the artificial entity. Moreover, in real life humans also incorporate a variety of social roles and different identities. In consequence, it is not fruitful to create "the" perfect persona, but instead to provide the user with different opportunities to attribute roles and personality. We have to go beyond imitation of single human roles toward a genuine companion identity,

which might be a collection of different identities.

3   Methods for artificial companion research

Since research in the domain of companion technologies is often interdisciplinary, a lot of different research methods have been applied. We argue that different methodologies need to be combined and in general advise to follow a multi-method approach (Ganster et al. 2010; von der Pütten et al. 2011b). First, multi-methodology compensates for the limitations every method entails. Within the combination of self-reported data and objectively obtained data, the latter can dispel doubts whether the self-reported data is affected by demand characteristics or socially desirable behaviour. Conversely, self-report often offers more possibilities to interpret the objectively obtained data. To give an example, in a study by Rosenthal-von der Pütten et al. (2013), investigating participants´ emotional reactions during videos showing a robot in a friendly or violent interaction with a human, self-reported data on the emotional state of the participants and psychophysiology measures (skin conductance and heart rate) were assessed. Participants indicated to feel more negatively after the reception of the video showing the robot being maltreated by the human. Moreover, they showed higher levels of physiological arousal. In combing these methods, the physiological arousal could be interpreted as increased negative response. Considering the higher physiological arousal, it seems very unlikely that the differences in the self-reported emotional states were due to socially desirable behaviour. And second, different methods yield different findings, because they address different aspects of human-artefact interaction. For instance, within the EU project SERA (www.sera-project.eu) diverse methods were used to examine human ro-

bot long-term relationships ranging from quantitative analysis of verbal and nonverbal behaviour (e.g., speech, eye-contact, smiling) during interaction, to post-hoc semi-structured interviews on usability, personal experience and relationship building (both reported in von der Pütten et al. 2011b) and case-based Conversation Analysis (Payr 2010). In this set-up, elderly healthy participants were interacting with a rabbit shaped robot which served as an advisor for physical activity. The system was installed in the participants´ homes for three consequent iterations of data collection, each lasting approximately ten days. The quantitative analysis of behaviour revealed that people spoke to the robot and showed nonverbal behaviour although the robot was not able to perceive this behaviour, which was known to the participants. The behaviour towards the robot as well as behaviour change over time was foremost idiosyncratic. From the interviews we were able to identify certain types of users. Users experienced with health-related technology regarded the robot more as a technology with the purpose to assist them in daily tasks, while others valued the social aspect of the robot. The latter group of users gave the robot a name and stated to miss the rabbit when it was gone. The Conversation Analysis of diverse interaction of one of the participants revealed that the participant treated the rabbit in very different ways depending on whether the participant was alone or in the presence of a third person (Payr 2010). In sum, the various methods delivered results with regard to participants' verbal and nonverbal behaviour (quantitative analysis), user types (interviews) and with regard to the question how individual users integrate the artefact into daily social interactions with others. Only the combination of these very different methods allowed a comprehensive examination of human-robot relationship building. It led to a deep understanding of what

was going on and allowed for the identification of issues worth to be investigated in more detail in the future.

Although the idea of companion technologies is to incorporate a certain role and take over certain tasks over a longer period of time, long-term studies are still scarce. There also is a lack of field studies with regard to companion technologies. Both are, however, necessary to investigate how long-term relationships are established (von der Pütten et al. 2011b).

In the following, we want to present diverse methodologies with regard to how they are used in HRI today and what additional potential they have not exploited so far. Methodological instruments can be differentiated between subjectively measurable aspects on the one hand and objectively measurable aspects or behavioural data, respectively, on the other hand.

3.1 Subjective measures

Subjective measures are commonly used in psychological research and include self-report via questionnaires and interviews. In human-artefact interaction research, scales address, e.g., socio-emotional aspects of the interaction or an evaluation of the agent/robot itself. For this purpose, on the one hand, standard instruments from social psychology are used to cover different aspects such as stereotypes and person perception. For instance, the Positive and Negative Affect Scale (PANAS, Watson et al. 1988) is often used when emotional experiences are evaluated (e.g., Rosenthal-von der Pütten et al. 2013; von der Pütten et al. 2008). On the other hand, some scales were especially created for use in human-agent/robot interaction studies, such as the Agent Persona Instrument (API) by Baylor and Ryu (2003) and the Attitude Towards Agents Scale (ATAS) (van Eck & Adcock 2003). Other scales were designed to be used across different media/technologies, e.g., questionnaires on immersion, physical and

social presence (e.g., Biocca & Harms 2002; Lombard et al.).

There are techniques and scales that allow for an evaluation of more application oriented aspects like appearance (e.g., card sort assignments; Cowell & Stanney 2003), perceived efficiency (e.g., Krämer & Nitschke 2002), believability and trust in a system (e.g., Sproull et al. 1996). Besides questionnaires, also interviews are frequently used in human-artefact interaction studies to shed light on diverse topics of interest, giving researchers the opportunity to gain a deeper understanding of participants' thoughts, opinions and attitudes (e.g., with regard to relationship building: Klamer & Ben Allouch 2010). In addition, less frequently used, yet informative methods exist. For instance, user diaries were used within the EU project LIREC where participants were provided with a Pleo for several weeks and were instructed to post their experiences with it in a blog.

And finally, to investigate the influence of personality traits in HRI/HAI a lot of standardized questionnaires can be adapted or employed "as are" in human-agent/robot interaction studies. Indeed, participants' personality traits (such as agreeableness, extraversion, shyness) have been shown to have great influence on the evaluation of artificial entities, on participants' emotional experiences, and their actual behaviour during the interaction (e.g., von der Pütten et al. 2010). Relatively new are instruments measuring personality traits directly connected to agents or robots, like the Robot Anxiety questionnaire (Nomura et al. 2007) or the Negative Attitudes Towards Robots questionnaire (Nomura et al. 2006), which have been also shown to be influential.

3.2 Objective measures

Investigations in HRI and HAI use diverse objective measures, ranging from conventional audio and video

analysis, to eye-tracking, psychophysiology and fMRI.

Many researchers make use of natural language recordings to be able to identify certain characteristics of the participant´s use of language and changes occurring during the interaction with the robot/agent. Language parameters may for example be the number and/or length of the user´s utterances (von der Pütten et al. 2011c), the number of overlapping speech and hesitations, the percentage of pause fillers, prolonged words and incomplete words compared to the total number of words (e.g., Gratch et al. 2007). Especially in natural language analysis, qualitative analyses can and should go hand in hand with quantitative analyses (e.g., analysis of intimacy of answers: von der Pütten et al. 2011c; discourse analysis: Payr 2010).

The analysis of video recordings is also widely used. Here, especially nonverbal cues are of interest. As in robot and agent research subjects' nonverbal behaviour during interactions with the robot can provide useful information. Video recordings are used here as well, showing, for instance, that participants mimic an agent´s nonverbal behaviour (Krämer et al. 2013), apply situationally appropriate nonverbal behaviour like waving while saying goodbye (von der Pütten et al. 2009), and display socio-emotional nonverbal behaviour (von der Pütten et al. 2011b).

In the context of studying human-robot/agent interaction, eye tracking may be a useful tool for evaluating artificial entities, because eye tracking gives information about where participants look at and for how long. Moreover, eye tracking can be used to find out whether a subject shows the same behaviour towards a robot or agent as he would show towards a human being (e.g., MacDorman et al. 2005; Shimada et al. 2010).

Also psychophysiology (e.g., electrodermal activity (EDA), electrocardiograms (ECG) and electroencephalograms (EEG)) can provide information not only as a medical means to monitor a patient's condition, but also to address psychological research questions. With regard to robots and agents, the data can be used to gain information about the participant's reactions towards the robot or agent. When measured during interaction with a robot or agent, EDA or ECG data might provide information about the subject's arousal and indicate stressful experiences in the encounter with the robot/agents (e.g., Rosenthal-von der Pütten et al. 2013; Bethel et al. 2007). This method is, however, not widely used in HRI studies.

Relatively new to HAI and HRI research, but of increasing popularity, is the use of functional magnetic resonance imaging. Studies utilizing fMRI address diverse research questions: Do robot and human stimuli result in similar brain activation with regard to movement (Chaminade & Cheng 2009), emotional expression (Chaminade et al. 2010), Theory of Mind (Frank et al. 2008), empathy with others (von der Pütten et al. 2011a), etc.

4 Conclusion

The aim of this paper was to provide a summary of the state-of the-art for research on companions from a social psychology perspective with regard to theoretical and methodological issues. In this line, we summarized psychological theories on sociability in human-human interaction and discussed the applicability of these assumptions on the sociability of artefacts. Sociability is obviously a complex concept which we tried to disentangle by introducing three levels of sociability: the actual communication, the relationship, and the roles that might be assigned. If we would like to

provide sociability in its complexity, we have to attend to all three levels.

With regard to the actual communication (level one) it can be concluded that there is no real alternative to utilizing human-human interaction theories. This is due to the fact that humans in their interactions with robots and agents will not stop to employ and expect the communicative mechanisms they are used to (e.g., perspective taking, common ground, Theory of Mind). Although, Theory of Mind is now regarded as fruitful concept that should be implemented (see Breazeal et al. 2004; Peters 2006; Marsella & Pynadath D.V 2005), there are only few attempts to actually model and implement ToM-like abilities, also due to the complexity of ToM capabilities. Thus, Krämer et al. (2011) introduced a categorization of ToM capabilities in order to simplify realization. Moreover, we presented an alternative to the model of human-human communication: human-dog communication. Although one might initially think that implementing interactions referring to human-dog communication is easier, it has been shown that human-dog communication largely relies on the same mechanisms as human-human communication (e.g., joint attention; Miklósi 2009), because dogs have been adapted to the human communication system by natural and breed selection (Tomasello 2008).

When it comes to relationship building (level two) the conclusion is more complex. On the one hand it makes sense to draw on some of the HHI theories presented here and use their "benefits". Developers, for instance, should design physically attractive agents and robots. Moreover, reciprocal liking can be easily exploited to foster relationship building. On the other hand, we saw from diverse (long-term) field studies, that some users incorporate companion technologies into their lives differently. Some form an emotional relationship, some

treat those devices as the piece of technology they are. Thus, it is questionable whether HHI relationship theories, like the social exchange theory, are applicable for HRI/HAI, i.e. whether humans evaluate human-artefact relationships similarly to human-human relationships. Moreover, it can be debated whether this is desirable. In conclusion, although it is difficult to establish a radically different model for human-robot/agent interaction, we would not say that merely human-human communication should be used as a framework for companions. Since there is little empirical work on human-artefact relationships, there is also little known on the nature of these relationships. Therefore, more long-term studies and field studies are needed.

It also can be debated whether companions have to assume a role modelled after human roles (level three) or whether new role models for companions can be established. Robots and agents are devices that satisfy certain needs of their owners and have their uses and functions in the owners' lives. Empirical studies have shown that people integrated these devices (e.g., robotic pets: Fernaeus et al. 2010; Joana Dimas et al. 2010; and robot vacuum cleaners: Sung et al. 2010; Forlizzi 2007) into their lives. When companions have the function to support the owners' health, well-being, and independent living, however, they adopt a role that goes far beyond that of a vacuum cleaner, and they have to be able to maintain that role over a longer period of time.

Thus, long-term field studies are necessary to investigate how long-term relationships are built and re-built on the micro-level of conversational interaction. In our pleading for the importance of multi-methodological research we stressed that future research should also include qualitative aspects, since it was shown that qualitative analyses were especially help-

ful for observing and understanding people's idiosyncratic reactions (e.g., in the SERA project, see von der Pütten et al. 2011b; Payr 2010).

Altogether, we introduced different levels of sociability and the corresponding theories in human-human communication. We pointed out which theories and concepts we regard as mandatory (e.g., perspective taking, common ground, Theory of Mind), useful (e.g., attractiveness, reciprocal liking) or marginally useful (e.g., social exchange theory, human role models) or limited in their explanatory power, respectively. Moreover, we summarized the state-of-the-art and emphasized the research gaps with regard to long-term field studies and on a theoretical level with regard to Theory-of-Mind- like abilities in robots. And finally, we emphasized that working on companion technologies (theoretically and technologically) without considering the human user and his/her needs, perceptions, and communication patterns will not be useful.

Lt. Cmdr. Data: Jenna – are we no longer... a couple?
Lt. Jenna D'Sora: No, we're not.
Lt. Cmdr. Data: Then I will delete the appropriate program.

~THE END~

References

Angeli, Antonella de, et al., 2006: Misuse and abuse of interactive technologies. In: *CHI´06: Proceedings of the Conference on Computer-Human Interaction*, Montréal, Québec, Canada, 1647–1650.

Aronson, Elliot/Timothy D. Wilson/Robin M. Akert, 2010: *Social psychology*, 7th edn. Prentice Hall, Upper Saddle River, NJ.

Baumeister, Roy F./Marc R. Leary, 1995: The need to belong: Desire for interpersonal attachments as a fundamental human motivation. In: *Psychological Bulletin* 117, 3, 497–529.

Baylor, Amy L./Jeeheon Ryu, 2003: The API (Agent Persona Instrument) for Assessing Pedagogical Agent Persona. In: David Lassner/Carmel McNaught (eds.) *World Conference on Educational Multimedia, Hypermedia and Telecommunications 2003*. AACE, Honolulu, Hawaii, USA, 448–451.

Baylor, Amy L./Soyoung Kim, 2008: The Effects of Agent Nonverbal Communication on Procedural and Attitudinal Learning Outcomes. In: Helmut Prendinger/James C. Lester/Mitsuru Ishizuka (eds.), *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer Berlin Heidelberg, 208–214.

Bell, Linda/Joakim Gustafson/Mattias Heldner, 2003: Prosodic adaptation in human–computer interaction. In: Maria-Josep Solé/Daniel Recasens i Vives (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*. Universitat Autónoma de Barcelona, Barcelona, 2453–2456.

Benninghoff, Brenda, et al., 2012: Theory of Mind in Human-Robot-Communication: Appreciated or not? In: Annette Kluge/Regina Söffker (eds.), *2. Interdisziplinärer Workshop Kognitive Systeme: Mensch, Teams, Systeme und Automaten*.

Berscheid, Ellen/E. Walster, 1978: *Interpersonal attraction*. Reading, MA: Addison-Wesley.

Bethel, Cindy L, et al., 2007: Psychophysiological experimental design for use in human-robot interaction studies. In: *Collaborative Technologies and Systems, 2007. CTS 2007. International Symposium on Collaborative Technologies and Systems, 2007. CTS 2007*, 99–105.

Biocca, Frank/Chad Harms, 2002: Defining and measuring social presence: Contribution to the net-worked minds theory and measure. In: Felix R. Gouveia/Frank Biocca (eds.), *Presence 2002. 5th Annual International Workshop on Presence*, 7–36.

Breazeal, Cynthia L., 2002: *Designing Sociable Robots*. Cambridge, MA: MIT Press.

Breazeal, Cynthia, et al., 2004: Tutelage and collaboration for humanoid robots. In: *International Journal of Humanoid Robotics* 01, 2, 315–348.

Breazeal, Cynthia/Jesse Gray/Matt Berin, 2011: Mindreading as a Foundational Skill for Socially Intelligent Robots. In: Makoto Kaneko/Yoshihiko Nakamura (eds.), *Robotics Research*. Berlin, Heidelberg: Springer Berlin Heidelberg, 383-394.

Cacioppo, John T./William Patrick, 2008: *Loneliness: human nature and the need for social connection*. New York: W. W. Norton and Company.

Carruthers, Peter/Peter K. Smith (eds.), 1996: *Theories of theories of mind*. Cambridge: Cambridge University Press.

Chaminade, Thierry/Gordon Cheng, 2009: Social cognitive neuroscience and humanoid robotics: Neurorobotics. In: *Journal of Physiology-Paris* 103, 3, 286–295.

Chaminade, Thierry, et al., 2010: Brain Response to a Humanoid Robot in Areas Implicated in the Perception of Human Emotional Gestures. *PLoS ONE* 5, 7, e11577 EP.

Clark, Herbert H., 1992: *Arenas of language use*. Chicago: University of Chicago Press.

Cowell, Andrew J./Kay M. Stanney, 2003: Embodiment and Interaction Guidelines for Designing Credible, Trustworthy Embodied Conversational Agents. In: *Proceedings of the 4th International Workshop on intelligent virtual agents. IVA 2003*. Berlin, London: Springer, 301–309.

Dautenhahn, Kerstin, 2004: Robots We Like to Live With? - A Developmental Perspective on a Personalized, Life-Long Robot Companion. In: *Proceedings of the 13th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2004)*, 17–22.

Dautenhahn, Kerstin/ Aude Billard, 1999: Bringing up robots or-the psychology of socially intelligent robots. In: Jeffrey Bradshaw/Jörg Muller/Oren Etzioni (eds.), *Proceedings of the third annual conference on Autonomous Agents. AGENTS '99*. New York: ACM Press, 366–367.

Dautenhahn, Kerstin, et al., 2005: What is a Robot companion - Friend, Assistant or Butler? In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005)*, 1488–1493.

Dion, Karen/Ellen Berscheid/Elain Walster, 1972: What is beautiful is good. In: *Journal of Personality and Social Psychology* 24, 3, 285–290.

Eimler, Sabrina C., et al., 2010: Following the white rabbit: a robot rabbit as vocabulary trainer for beginners of English. In: Gerhard Leitner/Martin Hitz/Andreas Holzinger (eds.), *HCI in work and learning, life and leisure: 6th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering, USAB 2010, Klagenfurt, Austria, November 4-5*. Berlin: Springer, 322-339.

Fernaeus, Yvla, et al., 2010: How do you play with a robotic toy animal? In: Narcís Parés (ed.), *IDC '10 Proceedings of the 9th International Conference on Interaction Design and Children*. New York: ACM, 39.

Fischer, Kerstin, 2006: *What computer talk is and isn't: Human-computer conversation as intercultural communication*. Saarbrücken: AQ-Verlag.

Forlizzi, Jodi, 2007: How robotic products become social products: an ethnographic study of cleaning in the home. In: Cynthia L. Breazeal, et al., (eds.), *HRI'07. Proceedings of the ACM/IEEE international conference on Human-robot interaction*. New York: ACM New York, 129–136.

Hegel, Frank, et al., 2008: Theory of mind (ToM) on robots: a functional neuroimaging study. In: *HRI'08. Proceedings of the 3rd ACM/IEEE international conference on Human Robot Interaction*. Amsterdam, The Netherlands: ACM, 335–342.

Frith, Uta/Chris Frith, 2003: Development and neurophysiology of mentalizing. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 358, 1431, 459–473.

Ganster, Tina, et al., 2010: Methodological Considerations for Long-Term Experience with Robots and Agents. In: Robert Trappl (ed.), *European Meetings on Cybernetics and Systems Research (EMCSR) 2010*, Vienna, Austria, 565–570.

Gold, Joel A./Richard M. Ryckman/Norman R. Mosley, 1984: Romantic Mood Induction and Attraction to a Dissimilar Other: Is Love Blind? In: *Personality and Social Psychology Bulletin* 10, 3, 358–368.

Gratch, Jonathan, et al., 2007: Can virtual humans be more engaging than real ones? In: Julie A. Jacko (ed.), *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer Berlin Heidelberg, 286–297.

Homans, George C., 1961: *Social behavior: Its elementary forms*. New York: Harcourt Brace.

Hüttenrauch, Helge, et al., 2004: Involving users in the design of a mobile office robot. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews* 34, 2, 113–124.

Ishiguro, Hiroshi, 2006: Interactive humanoids and androids as ideal interfaces for humans. In: Cécile L. Paris, et al., (eds.), *IUI '06. Proceedings of the 11th international conference on Intelligent User Interfaces*. New York: ACM Press, 2–9.

Dimas, Joana, et al., 2010: Pervasive Pleo: Long-term Attachment with Artificial Pets. In: *Please enjoy! Workshop on playful experiences in Mobile HCI*. Lisbon, Portugal: ACM.

Kahn, Peter H., et al., 2012: "Robovie, you'll have to go into the closet now": Children's social and moral relationships with a humanoid robot. In: *De-

velopmental Psychology* 48, 2, 303–314.

Kappas, Arvid, 2005: My happy vacuum cleaner. In: *ISRE General Meeting, Symposium on Artificial Emotions*, Bari.

Kelly, Harold H./John W. Thibaut, 1978: *Interpersonal relations: A theory of interdependence.* New York: Wiley.

Kenny, David A., 1994: Using the social relations model to understand relationships. In: Ralph Erber/ Robin Gilmour (eds.), *Theoretical frameworks for personal relationships.* Hillsdale, England: Lawrence Erlbaum Associates, Inc., 111–127.

Koay, Kheng Lee, et al., 2009: Five Weeks in the Robot House – Exploratory Human-Robot Interaction Trials in a Domestic Setting. In: *Second International Conferences on Advances in Computer-Human Interactions, 2009. ACHI '09,* 219–226.

Klamer, Tineke/Somaya Ben Allouch, 2010: Acceptance and Use of a Zoomorphic Robot in a Domestic Setting. In: Robert Trappl (ed.), *European Meetings on Cybernetics and Systems Research (EMCSR) 2010*, Vienna, Austria, 553–558.

Klatt, Jennifer/Stacy Marsella/Nicole C. Krämer, 2011: Negotiations in the Context of AIDS Prevention: An Agent-Based Model Using Theory of Mind. In: H. Hannes Vilhjálmsson, et al., (eds.), *Lecture Notes in Computer Science.* Berlin, Heidelberg: Springer Berlin Heidelberg, 209–215.

Krach, Sören, et al., 2008: Can Machines Think? Interaction and Perspective Taking with Robots Investigated via fMRI. *PLoS ONE* 3, 7, e2597.

Krämer, Nicole C., 2010: Psychological Research on Embodied Conversational Agents: The Case of Pedagogical Agents. In: *Journal of Media Psychology: Theories, Methods, and Applications* 22, 2, 47–51.

Krämer, Nicole C., et al., 2013: Smile and the world will smile with you - The effects of a virtual agent's smile on users' evaluation and behavior. In: *International Journal of Human-Computer Studies* 71, 335–349.

Krämer, Nicole C./Julia Nitschke, 2002: Ausgabemodalitäten im Vergleich: Verändern sie das Eingabeverhalten der Benutzer? In: Ruth Marzi (ed.) *Bedienen und Verstehen. 4. Berliner Werkstatt Mensch-Maschine-Systeme*. Düsseldorf: VDI-Verlag, 231–248.

Krämer, Nicole C./Gary Bente/Jens Piesk, 2003: The ghost in the machine. The influence of Embodied Conversational Agents on user expectations ans user behavior in a TV/VCR application. In:

Gerald Bieber/Thomas Kirste (eds.), *IMC Workshop 2003, Assistance, Mobility, Applications.* Stuttgart: Fraunhofer IRB Verlag.

Krämer, Nicole C., et al., 2011: Theory of Companions: What can Theoretical Models contribute to Applications and Understanding of Human-Robot Interaction? In: *Applied Artificial Intelligence* 25, 6, 474–502.

Krauss, Robert M./Susan R. Fussell, 1991: Perspective-taking in communication: Representations of others' knowledge in reference. In: *Social Cognition* 9, 1, 2–24.

Kubitschek, Warren N./Maureen T. Hallinan, 1998: Tracking and Students' Friendships. In: *Social Psychology Quarterly* 61, 1, 1–15.

Lester, James C., et al., 2000: Deictic and emotive communication in animated pedagogical agents. In: Justine Cassell (ed.) *Embodied conversational agents.* Cambridge: MIT Press, 123–154.

Lombard, Matthew, et al., 2000: Measuring presence: A literature-based approach to the development of a standardized paper-and-pencil instrument. In: *Presence 2000: The Third International Workshop on Presence.*

MacDorman, Karl F., et al., 2005: Assessing human likeness by eye contact in an android testbed. In: *Proceedings of the XXVII Annual Meeting of the Cognitive Science Society.*

Marsella, Stacy/David V. Pynadath, 2005: Modeling influence and theory of mind. Artificial Intelligence and the Simulation of Behavior. In: *Proceedings of the Joint Symposium on Virtual Social Agents: AISB'05: Social Intelligence and Interaction in Animals, Robots and Agents*, 199–206.

Matarić, Maja J., et al., 2007: Socially Assistive Robotics for Post-Stroke Rehabilitation. In: *Journal of NeuroEngineering and Rehabilitation* 4, 1, 5.

Miklósi, Ádam, 2009: Evolutionary approach to communication between humans and dogs. In: *Veterinary Research Communications* 33, S1, 53–59.

Nass, Clifford/Youngme Moon, 2000: Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues* 56, 1, 81–103.

Nickerson, Raymond S., 1999: How we know—and sometimes misjudge—what others know: Imputing one's own knowledge to others. In: *Psychological Bulletin* 125, 6, 737–759.

Nomura, Tatsuya, et al., 2006: Measurement of negative attitudes toward robots. In: *Interaction Studies* 7, 3, 437–454.

Nomura, Tatsuya, et al., 2007: Measurement of Anxiety toward Robots. In:

RO-MAN'07. Proceedings of the 14th IEEE International Symposium on Robot and Human Interactive Communication. 372–377.

Payr, Sabine, 2010: Ritual or Routine: Communication in long-term Relationships with Companions. In: Robert Trappl (ed.), *European Meetings on Cybernetics and Systems Research (EMCSR) 2010*, Vienna, Austria, 559–564.

Peters, C, 2006: A perceptually-based theory of mind for agent interaction initiation. In: *International Journal of Humanoid Robotics* 3, 3, 321–339.

Premack, David/Ann J. Premack, 1995: Origins of human social competence. In: Gazzaniga, M. S. (ed.) *The cognitive neurosciences.* Cambridge: MIT Press, 205–218.

Premack, David/Guy Woodruff, 1978: Does the chimpanzee have a theory of mind? In: *Behavioral and Brain Sciences* 1, 4, 515–526.

Reeves, Byron/Clifford Nass, 1996: *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places.* Cambridge: Cambridge University Press.

Robins, Ben, et al., 2012: Scenarios of robot-assisted play for children with cognitive and physical disabilities. In: *Interaction Studies* 13, 2, 189–234.

Rommeveit, Ragnar, 1974: *On message structure: A framework for the study of language and communication.* New York: Wiley.

Rosenthal-von der Pütten, Astrid M., et al., 2013: An Experimental Study on Emotional Reactions towards a Robot. In: *International Journal of Social Robotics* 5, 1, 17–34.

Ryan, Richard M./Edward L. Deci, 2000: The darker and brighter sides of human existence: Basic psychological needs as a unifying concept. In: *Psychological Inquiry* 11, 4, 319–338.

Scassellati, Brian, 2002: Theory of Mind for a Humanoid Robot. In: *Autonomous Robots* 12, 1, 13–24.

Shechtman, Nicole/Leonard Horowitz, 2003: Media inequality in conversation: how people behave differently when interacting with computers and people. In: *CHI'03. Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM, Ft. Lauderdale, Florida, USA, 281–288.

Shimada, Michihiro, et al., 2010: Effects of Observing Eye Contact between a Robot and Another Person. In: *International Journal of Social Robotics* 3, 2, 143–154.

Sobieraj, Sabrina, 2012: *What is virtually beautiful is good-Der Einfluss physiognomischer und nonverbaler Gesichtsmerkmale auf die Attribution von Attraktivität, sozialer Kompetenz und Dominanz.* PhD thesis, Duisburg.

Sproull, Lee, et al., 1996: When the Interface Is a Face. In: *Human-Computer Interaction* 11, 2, 97–124.

Sung, Ja-Young/Rebecca Grinter/Henrik I. Christensen, 2010: Domestic Robot Ecology. In: *International Journal of Social Robotics 2*, 4, 417–429.

Swann, William B./Alan Stein-Seroussi/Shawn McNulty, 1992: Outcasts in a white-lie society: The enigmatic worlds of people with negative self-conceptions. In: *Journal of Personality and Social Psychology* 62, 4, 618–624.

Syrdal, Dag S., et al., 2010: Video prototyping of dog-inspired non-verbal affective communication for an appearance constrained robot. In: Carlo A. Avizzano/Emanuele Ruffaldi (eds.), *Proceedings of the 19th IEEE International Symposium on Robot and Human Interactive Communication. Ro-Man 2010.* Piscataway: IEEE, 632–637.

Thibaut, John W./Harold H. Kelley, 1959: *The social psychology of groups.* New York: Wiley.

Tomasello, Michael, 2008: *Origins of human communication.* Cambridge: MIT Press.

van Eck, Richard/Amy Adcock, 2003: Reliability and factor structure of the Attitude Toward Agent Scale (ATAS). In: *Paper presented at the annual meeting of the American Educational Research Association 2003.*

von der Pütten, Astrid M./Nicole C. Krämer, 2012: A survey on robot appearances. In: *Proceedings of the 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI`12)*, 267–268.

von der Pütten, Astrid M., et al., 2011a: I not robot! Empathetic reactions towards robots. In: *Poster presented at the Annual Meeting of the International Society for Research on Emotion, Kyoto, Japan.*

von der Pütten, Astrid M./Sabrina C. Eimler/Nicole C. Krämer, 2011b: Living with a Robot Companion: Empirical Study on the Interaction with an Artificial Health Advisor. In: *ICMI'11: Proceedings of the 2011 ACM International Conference on Multimodal Interaction.* New York: ACM, 327.

von der Pütten, Astrid M., et al., 2011c: Quid Pro Quo? Reciprocal Self-disclosure and Communicative Accomodation Towards a Virtual Interviewer. In: H. Hannes Vilhjálmsson, et al. (eds.), *Lecture Notes in Computer Science.* Berlin, Heidelberg: Springer Berlin Heidelberg, 183–194.

von der Pütten, Astrid M./Nicole C. Krämer/Jonathan Gratch, 2010: How our personality shapes our interactions with virtual characters: implications for research and development. In: Jan Allbeck, et al. (eds.), *Intelligent Virtual Agents 2010*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 208–221.

von der Pütten, Astrid M., et al., 2008: Comparing Emotional vs. Envelope Feedback for ECAs. In: *Proceedings of the 8th international conference on Intelligent Virtual Agents*. Springer-Verlag, Tokyo, Japan, 550–551.

von der Pütten, Astrid M., et al., 2009: The Impact of Different Embodied Agent-Feedback on Users' Behavior. In: Zofia Ruttkay, et al. (eds.) *Intelligent Virtual Agents 2009*. Springer-Verlag, Amsterdam, The Netherlands, 549–551.

Walker, Marilyn A., et al., 2002: Automatically Training a Problematic Dialogue Predictor for a Spoken Dialogue System. In: *Journal of Artificial Intelligence Research*, 293–319.

Watson, David/Auke Tellegen/Lee A. Clark, 1988: Development and Validation of Brief Measure of Positive ans Negative Affect: The PANAS Scales. In: *Journal of Personality and Social Psychology* 54, 6, 1063–1070.

Watzlawick, Paul/Janet Beavin/Don Jackson, 1967: *Pragmatics of Human Communication. A study of interactional patterns, pathologies, and paradoxes.* New York: Norton.

Waytz, Adam, et al., 2010: Causes and consequences of mind perception. In: *Trends in Cognitive Sciences* 14, 8, 383–388.

Whiten, Aandrew (ed.), 1991: *Natural Theories of Mind: Evolution, Development and Simulation of Everyday Mindreading*. Oxford: Basil Blackwell.

Yaghoubzadeh, Ramin, 2011: FamCHAI: An Adaptive Calendar Dialogue System. In: Joseph Konstan (eds.) LNCS. Berlin-Heidelberg: Springer Berlin-Heidelberg, 458-461.