MASTER THESIS

Expectations and Human-Robot Interaction

The influence of robot expectations on personality attribution, impressions and anthropomorphism

Jarno Rutjes

S1020250

UNIVERSITY OF TWENTE
ENSCHEDE
GW-CMC

EXAMINATION COMMITTEE

Dr. S. Ben Allouch, GW-MCO

MSc M.M.A. de Graaf, GW-MCO

VERSION
1.4

UNIVERSITY OF TWENTE.

02/07/2013

Expectations and Human-Robot Interaction; The influence of robot expectations on personality attribution, impressions and anthropomorphism

Samenvatting

Tegenwoordig worden steeds meer sociale robots op de consumentenmarkt gebracht. Het is daarom belangrijk om te onderzoeken hoe en waarom mensen interactie voeren met een sociale robot. Binnen deze studie wordt de volgende definitie voor een sociale robot gehanteerd: Een sociale robot is een autonome of semiautonome robot, welke interacteert en communiceert met mensen en rekening houdt met de gedragsnormen die mensen hanteren. Deze studie onderzoekt mens-robot interactie en de invloed van verwachtingen die mensen hebben van deze robot op impressies van de sociale robot, de mate waarin men de robot anthropomorphiseerde en of men geneigd was eigen persoonlijkheidseigenschappen te reflecteren op een sociale robot. Ook is er onderzocht wat de invloed is van een overeenkomende persoonlijkheid tussen mens en robot op de impressies van deze robot.

Voor het uitvoeren van dit onderzoek zijn de verwachtingen van de participanten gemanipuleerd. Hierbij is onderscheid gemaakt tussen twee condities, één conditie waarbij men lage verwachtingen had van de sociale robot en één conditie waarbij men hoge verwachtingen had van de sociale robot. Voor het toetsen van de hypothesen zijn er verschillende meetinstrumenten gebruikt, deze waren: Saucier's (1994) "Big-Five inventory set", Bartneck's (2009) antropomorfisme items, Kanda's (2001) tegenoverstellende paren voor het meten van impressies en de "Personal Innovativeness In Technology" schaal (Agarwal & Prasad, 1998).

De resultaten gaven aan dat participanten meer positieve impressies hadden van een sociale robot, wanneer er een overeenkomende gewetensvolle persoonlijkheid was. Deze bevinding sluit zich aan bij de "similarity-attraction" hypothese (Byrne et al, 1986). Participanten die zichzelf een aangename persoonlijkheid toe wezen, waren geneigd deze persoonlijkheidseigenschap ook toe te wijzen aan een sociale robot. Dit resultaat komt overeen met de "assumed similarity" en "attributive projection" theorie (Cronbach, 1995; Holmes, 1978). Bovendien, was dit resultaat het meest zichtbaar in de lage verwachtingen conditie. De resultaten met betrekking tot lage of hoge verwachtingen van de robot indiceerden dat mensen met hoge verwachtingen meer positievere impressies hadden van een sociale robot dan mensen met lage verwachtingen. Dit betekend dat de "Self-Fulfilling Prophecy" en de "Confirmation Bias", in geval van hoge robot verwachtingen, binnen dit onderzoek bevestigt worden (Merton, 1948; Nickerson, 1998). Ook bleken mensen met hoge verwachtingen de sociale robot in hogere mate te anthropomorphiseren dan mensen met lage verwachtingen. Op basis van de resultaten recommenderen wij dat toekomstig onderzoek zich verder zal moeten toespitsen op het effect van verwachtingen op mens-robot interactie, waarbij gebruik gemaakt zal worden van verschillende typen robots.

Abstract

More and more social robots are introduced to the consumer market. This means that it is important to investigate how and why humans interact with a social robot. The current study investigated if expectations of a social robot would influence the impressions humans had of a social robot, the degree to which humans anthropomorphised a social robot and to what extent humans would assign their own personality traits to that of the social robot. We also studied the influence of matching personality traits, between human and robot, on the impressions that this human had of the social robot. The results, from a between-participants experiment (high expectations and low expectations), indicated that participants had more positive impressions of a social robot when there was a matching conscientiousness personality. This finding supports the similarity-attraction hypothesis (Byrne et al, 1986). The assumed similarity and attributive projection theory were supported within this study for the agreeableness personality trait (Cronbach, 1995; Holmes, 1978). Participants who described themselves as being agreeable, tended to assign this personality trait to a social robot. This result was most visible in the low expectations condition. We also found that participants within the high expectations condition had more positive impressions of a social robot in comparison with participants within the low expectations conditions. This means that the Self-Fulfilling Prophecy and the Confirmation Bias, in case of high expectations, were confirmed within this study (Merton, 1948; Nickerson, 1998). We also found similar results for anthropomorphism. Participants within the high expectations condition anthropomorphised the social robot to a higher degree than participants within the low expectations condition. We recommend that future research should also study expectations and how it influences human-robot interaction, whereby they use a different kind and/or type of robot.

Introduction

Humans often interact with different objects, especially with their computers, smartphones and other digital objects. It has been confirmed that research in the field of human-computer interaction (HCI) can be applied to an area of increasing interest, namely that of human-robot interaction (Lee et al., 2005). The way humans interact with a robot is called Human-Robot Interaction (HRI). Robots, especially social robots, are more and more entering our homes, whereby these robots give us new interaction opportunities. Bartlett et al. (2004) have predicted that social robots will be as widespread as personal computers are used today. Therefore, a more recent and growing research area arises for investigating means by which humans can interact with a social robot. This study will focus on a specific type of robot, namely socially interactive robots. A socially interactive robot, which has the primary function to interact socially with humans, is an autonomous robot that interacts and communicates with humans by following social behaviours and rules attached to its role (Fong et al., 2003).

The personality of such a socially interactive robot is an important research domain for robot designers. Research in the past has shown that people tend to assign personality attributes to robots, this means that humans respond towards these robots in a natural and social manner (Lee et al., 2006). This study investigates if the attribution of personality traits to a socially interactive robot is affected by the personality traits of the user himself. We also investigate if this attribution of personality traits, the impressions of a social robot and anthropomorphising a social robot depends on the expectations we have for the interaction with a social robot.

In the following parts we present related work and theories in the area of socially interactive robots, human-robot interaction, attribution of personality traits and expectation setting.

Theoretical framework

Socially Interactive Robots

Social robots can be used for different purposes, such as a research platform, as a toy, using it for education or for therapeutic aids (Fong et al, 2003). Breazeal (2003) has divided social robots into four categories, namely: socially evocative, social interface, socially receptive and sociable robots. Socially evocative robots are designed to encourage humans to anthropomorphise the robot and they depend on the affective responses from the users. A social interface robot is characterized by recognition and manifestation of human-interaction modalities, like gestures and speech. Robots which are able to possess certain levels of social cognition, such as imitation and learning, can be described as socially receptive robots. Sociable robots have the capability to show high levels of social cognition, through which they can proactively seek social interaction and replicate human goals and desires.

Fong et al. (2003) added three more classes, namely socially situated, socially embedded and socially intelligent robots. Socially situated robots are surrounded by a social environment, which they perceive and react to. Socially embedded robots interact with humans within a social environment, are coupled with their social environment and they are aware (at least partially) of human interactional structures such as turn-taking. Socially intelligent robots have the capability of showing some aspects of human style social intelligence, which is based on human cognition and social competence. In general, we assume that a social robot must evoke meaningful social interaction with humans, which also actually elicit types of social responses.

The definition of a social robot within the context of this research is: A social robot is an autonomous or semi-autonomous robot that interacts and communicates with humans by following the behavioural norms expected by the people with whom the robot is intended to interact (Bartneck & Forlizzi, 2004).

Social responses towards a robot include affections and bonding with the robot, ontological perceptions of the robot as a social actor and performing social rules which are expected and applicable in a specific situation and environment (Lee et al., 2005). There are three key dimensions for a social robot that are considered as the critical factors for meaningful social interaction (Lee et al., 2005). First, anthropomorphic/zoomorphic forms and behaviours: Employing anthropomorphic/zoomorphic qualities in forms and behaviours, such as human/animal faces, human/animal bodies, speech recognition, gestures and touch reactions. Second, emotion: When a robot is capable of recognizing different emotional states and the robot himself can manifest emotion-like states, then it is likely that humans will show natural social responses to this robot. This is likely, because the human ability to recognize and manifest emotion is one of the key factors that will determine successful interactions. Third, personality: Humans prefer to interact with a robot that is able to manifest a compelling personality. It sounds obvious that humans experience a qualitatively better human-robot interaction, when the robot has a friendly personality than a nonfriendly robot personality (Scheeff et al., 2002). A robot can express his personality by emotional responses, physical attributes (e.g., size, shape and colour), motions, and the style of communication (Isbister & Nass, 2000).

Zoomorphic Robots

One specific type of robots that can manifest personality are zoomorphic robots. The number of social robots that has a zoomorphic embodiment is increasing. Robots with a zoomorphic embodiment are designed to imitate living creatures that can establish a human-robot relationship and they exhibit characteristics that are associated with domesticated animals. The human-robot relationship, in case of a zoomorphic embodiment, is in many ways similar to that of an owner-pet relationship (Kerepesi et al., 2006). This phenomenon will be explained in one of the following paragraphs. The 'uncanny valley effect' is a hypothesis which predicts that robots with human-like characteristics look and act almost, but not perfectly, like actual human beings, which causes a response of revulsion among human observers (Mori, 1970). Studying a robot with a zoomorphic embodiment has the advantage that the 'uncanny valley effect' can be better avoided, because a relationship between a human and a pet is simpler than a relationship between humans (Fong et al., 2003). Thus, it would be desirable to use a social robot with a zoomorphic embodiment within this study, because than we can better avoid the uncanny valley effect.

Human-Robot Interaction

Nowadays many robots are designed to interact with humans. This often requires that the robot can act autonomously. These autonomous robots can function as physical aids for elderly people, as museum-guide robots, as educational instruments, as therapeutic tools or as an entertainment robot which interacts socially with humans (Kerepesi et al., 2006). Research on how we interact with these robots is relatively new, but it has been confirmed that research in the field of human-computer interaction (HCI) can be applied to human-robot interaction (HRI) (Lee et al., 2005). Lee et al. (2006) tested if the "Computers are Social Actors" (CASA) research paradigm, which holds that humans apply wide sets of social characteristics to a computer (Nass & Moon, 2000), also could be applied to human-robot interaction. They found that the CASA research paradigm could be applied to the field of human-robot interaction. This means that humans often tend to respond socially towards these social robots, attribute personality to them and use personality-based social rules in their evaluation of the social robot for building and maintaining a relationship with this robot (Lee et al., 2005). Dautenhahn (2007) stated that the nature of human-robot interaction is related to human-human interaction. Therefore a possible reason why we treat a robot as if it is a real living creature, is that human cognitive development is socially situated and it is in the very nature of humans to be social (Dautenhahn, 2007). In order to perform the need of being social in interactions with social robots, we anthropomorphise these robots as if it has lifelike qualities and we interpret the robot's behaviour as being intentional (Breazeal, 2002). Based on the CASA research paradigm, we assume that gathering knowledge on how humans attribute lifelike qualities to a social robot, could result in a qualitatively better human-robot interaction. Therefore, it is important to know how anthropomorphism works and how a social robot can be even more anthropomorphised.

Anthropomorphism

Anthropomorphism means that humans attribute humanlike properties, characteristics, or mental states to real or imagined nonhuman objects, such as social robots (Epley, Waytz & Cacioppo, 2007). The extent to which humans anthropomorphise is determined by the Three-Factor-Theory of Anthropomorphism (Epley et al., 2007): (a) Elicited Agent Knowledge: The knowledge about humans in general or self-knowledge functions as a basis for induction, because this is already acquired knowledge and therefore this knowledge is more richly detailed than knowledge about nonhuman objects. This knowledge about human characteristics or self-knowledge is assumed to be more readily accessible at the time of judgment. (b) Effectance Motivation: Effectance describes the urge for humans to interact effectively with nonhuman objects in a specific environment. Humans tend to reduce the experience of uncertainty, when they are confronted with nonhuman objects, by attributing human characteristics and motivations to these nonhuman objects. This attribution increases the ability to make sense of the actions from certain nonhuman objects (e.g. robots) in the present and predicts the behaviour of these objects in the future. (c) Sociality Motivation: Sociality describes the need to maintain and establish social connections with other human beings. When humans experience absence of this social connection, then they will anthropomorphise to a higher degree in order to satisfy their motivation for social connection. These three factors of anthropomorphism can be used as guidelines for improving the interaction between a human and a social robot, because anthropomorphism helps humans to fulfil their need of being social in human-robot interaction (Breazeal, 2002).

Beyond Anthropomorphism

Sometimes the interaction with a social robot goes even further than anthropomorphising, whereby humans actually treat their social robots as if it is a living creature. Friedman et al. (2003) wanted to demonstrate this phenomenon, by suggesting that the relationship between people and their AIBO (a robotic dog) can be compared with the relationship that humans have with real living dogs. In addition, Bartlett et al. (2004) found that children actually treat AIBO as a living dog, where they make use of terms like 'he' or 'she' instead of 'it'. Beck et al. (2004) interviewed older adults and their relationship with AIBO. They found that these adults regarded AIBO as if it was a family member and some animal characteristics were attributed to the robot. Kerepesi et al. (2006) also studied the way humans (children and adults) interact with a social robot (AIBO) and a living dog puppy. They found that humans interact with the AIBO in some ways as if it were a living dog puppy and that the interaction with an AIBO is far more complex than a simple toy interaction. The fact that humans sometimes treat a social robot as a living creature means that they possibly will assign specific personality traits to a social robot. Therefore, it is important to investigate how personality within social robots can elicit meaningful social interaction.

Personality in Social Robots

It is possible, that a robot's personality leads to the fact that we often treat social robots as real living creatures. The personality of a robot is considered as one of the critical factors for meaningful social interaction. Humans often apply social norms and express social behaviour towards robots, where they treat robots as social entities (Lee et al., 2006). It is given that humans naturally pick up the personality of a robot from its design characteristics (Syrdal et al., 2007). This perceived personality invokes specific emotions in humans that are interacting with the robot (Hwang et al., 2012). This means that it is important to know what kind of personality a robot should embody when it is interacting with humans, because the perceived personality of the robot can influence the emotional responses of humans towards this robot.

Gathering knowledge about a robot's personality and the influence it has on how we interact with a social robot is important for understanding why we often treat robots as living creatures. This knowledge can also be used for improving the human-robot interactions, because personality is a key determinant in human social interactions (Tapus et al., 2008). Therefore, developing a clear, consistent and appealing robot personality is of major importance for meaningful human-robot interaction. The definition of personality for humans within the context of this study, based on the literature, is defined as: "Personality is a collection of individual differences, dispositions and temperaments that have consistency across situations and time" (Dryer, 1999, p. 274). We assume that this definition of personality can also be applied to social robots, because robots can manifest individual differences, dispositions and temperaments which can be consistent.

Personality Matching

How someone responds towards a specific personality of a robot depends on the individuals own personality. Some studies argued that the personality of a robot should match with that of the human user (Nakajima et al., 2003). Other studies indicated that the personality of a robot should match with his design purpose (Woods, 2006). More specific, Goetz and Kiesler (2002) found that humans find more enjoyment in the interaction with a happy robot, but they would rather comply with specific instructions when the robot had a more serious personality. This shows again that a robot's personality should match its design purpose. Woods et al. (2005) stated that there is also a possibility that humans try to match and project their own personality characteristics and styles to that of the robot with whom they are confronted with. In this case humans want to create and maintain an engaging interaction with the robot which is meaningful and gives them a familiar feeling. Woods et al. (2005) also proposed the opposite by stating that it is possible that humans do not want to perceive themselves as having the same personality of a robot, because they fear to lose their own identity. In this case they might show different personality traits or no personality traits, which may cause that humans do not want to bond with the specific robot (Woods et al., 2005).

There are also two opposite hypotheses in the field of matching personalities between humans: The similarity-attraction hypothesis and the complementarity principle (Byrne et al., 1986; Isbister & Nass, 2000). The similarity-attraction hypothesis tells that humans prefer to interact with humans that have the same personality traits. The complementarity principle holds that humans show complementary behaviours in their interpersonal interactions and prefer to interact with humans that have complementary personality traits (Isbister & Nass, 2000). Research in the field of human-computer interaction indicated that the similarity-attraction hypothesis holds true (Nass et al., 1995). Nass et al. (1995) compared interaction between dominant or submissive individuals with dominant or submissive computers. They found that individuals prefer to interact with a computer that was similar to themselves. The participants indicated that these interactions were more fun, more useful and more satisfying. However, Isbister and Nass (2000) found opposite results in their research on the personality of interactive characters (e.g. wizard characters within the Microsoft software). Contrary to the human-computer interaction literature, they found that humans prefer to interact with an interactive character that was complementary to them.

These contrary findings about personality matching also can be found in human-robot interaction studies. Lee et al. (2006) studied if the similarity-attraction hypothesis or the complementarity principle holds true in interaction between humans and the social robot AIBO. They found that individuals regarded the robot that had a complementary personality as being more intelligent, more attractive and more socially present than a robot that manifests a similar personality. Lee et al. (2006) also found critical evidence that humans actually responded to the social robot as it manifested personality, where they identified the personality and used personality-based social rules in their evaluation of the social robot. Bernier and Scassellati (2010) found evidence that the similarity-attraction hypothesis is the dominant force in human-robot interaction. They reported that participants rated a robot more positive when it displayed preferences similar to their own. However this specific study focused on attitudinal similarity instead of personality similarity. Park et al. (2012) found support that the similarity-attraction hypothesis held true in case of personality similarity. In this study, participants preferred to interact with a facial expression robot that had a similar personality to themselves. Because it seems that humans possess more positive impressions of a robot that is similar to themselves, the following hypothesis will be investigated in this study:

H1: Humans, according to the similarity-attraction hypothesis, have more positive impressions of a social robot when the robots' personality matches that of the human.

Assigning personality traits to a social robot

Humans often use personality as a social tool that helps them to explain and interpret behaviours of others. This is also the case in human-robot interaction, where humans do attribute personality traits to robots (Woods et al., 2005). Woods et al. (2005) also found evidence that younger people (18 to 30 years) attribute their own personality traits to a robot. The researchers declared that they attributed their own personality traits, because they wanted to have a better understanding of the interaction. On the other hand, older subject were less likely to attribute their own personality traits to that robot, which may be explained by the fact that older people attempt to keep their own identity separate from the robot's identity.

The tendency of perceiving an individual or object as having similar personality characteristics to their own can be described as assumed similarity (Cronbach, 1955). Assumed similarity refers to the belief that other people or social robots are similar to the self. Lee et al. (2009) stated that assumed similarity is a phenomenon that may occur when humans do not have valid trait-relevant information of the person/object they are interacting with. Srivastava et al. (2010) found evidence that the more a person described him- or herself as having a trait, the more likely the person was to see the trait in others. This means that humans tend to assign their own personality traits to the person they are in interaction with. Since there is little knowledge about this phenomenon in the context of human-robot interaction, we have formulated the following hypothesis:

H2: Humans, which have attributed particular personality traits to themselves, tend to assign similar personality traits to a social robot.

The Big Five

This study on matching and attribution of personality will use the Big Five inventory for measuring personality (Costa & McCrae, 1992). This inventory divides personality into five categories which represent broad areas of personality. These categories are: Extraversion-Introversion, Agreeableness, Conscientiousness, Emotional Stability and Openness.

The (a) Extraversion-Introversion category describes how outgoing and social a specific individual is. Extravert individuals are characterized as active, assertive, energetic, enthusiastic, outgoing and talkative. Introvert individuals have lower social engagement desires, have less need of stimulation and often feel the desire of being alone. (b) Agreeableness is about the tendency to find social harmony. Agreeable individuals are characterized as appreciative, forgiving, generous, kind, sympathetic and trusting. Disagreeable individuals are more egoistic and are generally unconcerned with someone else his well-being. (c) Conscientiousness individuals are well known for their planned behaviour instead of being more spontaneous. Humans that show high levels of conscientiousness are characterized as efficient, organized, planful, reliable, responsible and thorough. (d) The Emotional Stability category describes how individuals cope with negative emotions. Individuals that show low level of emotional stability are generally emotional unstable and are characterized as anxious, self-pitying, tense, touchy, unstable and worrying. (e) Openness means that a person is imaginative, independent minded and has divergent thinking. Characteristics of humans with a high level of openness are artistic, curious, imaginative, insightful, original wide interested (McCrae & John, 1992).

Goetz and Kiesler (2002) used the Big Five inventory for measuring personality of the robot user and the social robot. They investigated the influence of a robot displaying two different personalities on user compliance with an exercise routine. Their findings suggest that the Big Five inventory can be used for measuring personality within robots. Therefore, we will not only use the Big Five inventory for studying the attribution and possible matching of personality traits, we will also use the Big Five inventory for investigating the effect of specific expectations on the attribution of personality traits to a social robot.

Expectation Setting

Expectation setting could also influence the way humans assign personality traits to a social robot. An expectation is a strong belief that something will happen or will be the case (Oxford Dictionaries, n.d.). When these expectations are high and it turns out that these expectations cannot be fulfilled by a social robot, then people will become more disappointed in the social robot than people whose expectations were low (Paepcke & Takayama, 2010). This could possibly influence the way humans assign personality traits to a robot, because first impressions often shape the final appraisal we will have for a person or system (Rabin & Schrag, 1999). This is consistent with the Expectancy Violations theory (Burgoon & Hale, 1988). This theory attempts to explain one's reactions to unexpected behaviour of the person/object he/she is in interaction with. Expectancy Violations theory predicts that when expectations are not been kept and the interaction is unexpected, we will either reciprocate (match the behaviour) or compensate (acting opposite). The positive or negative judgement we make of the unexpected behaviour is called violation valence. Positive violations increase the attraction of the violator and negative violations decrease the attraction of the violator (Burgoon & Hale, 1988). So it is possible that when humans experience a negative unexpected behaviour of a social robot, they will be disappointed and will not accept the social robot as an interaction partner. Conversely it is possible that when humans experience a positive unexpected behaviour they will be surprised and will accept the social robot as an interaction partner.

The Self-Fulfilling Prophecy and the Confirmation Bias actually describe the opposite on outcomes of expectations. The Self-Fulfilling Prophecy describes how social beliefs, which can be characterized as expectations, influence how we interact with one another (Merton, 1948). This means that when humans beforehand do have high expectations of a robot's capabilities, they will be inclined to judge the robot as more capable then when expectations would set low. The Confirmation Bias can be related to this, because it also refers to the tendency for humans to seek or interpret evidence in a way that it will be similar to someone his beliefs or expectations (Nickerson, 1998).

It may be obvious that higher expectations could lead to the fact that humans are more likely to assign their own personality traits to a social robots in comparison with humans who have lower expectations, because we assume that humans with high expectations will see and treat the robot more as a living creature. This assumption that specific subjects will project their own traits and feelings onto others (e.g., fearful persons seeing others as fearful) is called attributive projection (Holmes, 1978). The theory of assumed similarity, as discussed before, connects to attributive projection. Humans may attribute their own personality traits onto others, because this enables him to see others as more familiar and less threatening (Halpern & Goldschmitt, 1976). Therefore, another argument arises which supports our assumption that humans with high expectations of a social robot are more likely to assign their own personality traits to a social robot. When humans have high expectations, they will have more motivations and reasons to search for familiar personality traits which makes the robot less threatening and therefore it would be more likely that projective attribution will occur for humans with high expectations in comparison with humans which have low expectations.

From the theories related to expectations, combined with projective attribution and impressions, we have formulated the following hypothesis:

H3: Humans who have high expectations of a social robot are more likely to assign their own personality traits to that of a social robot in comparison with humans that have low expectations for a social robot.

H4: Humans which have high expectations of a social robot will have more positive impressions of a social robot than humans which have low expectations of a social robot.

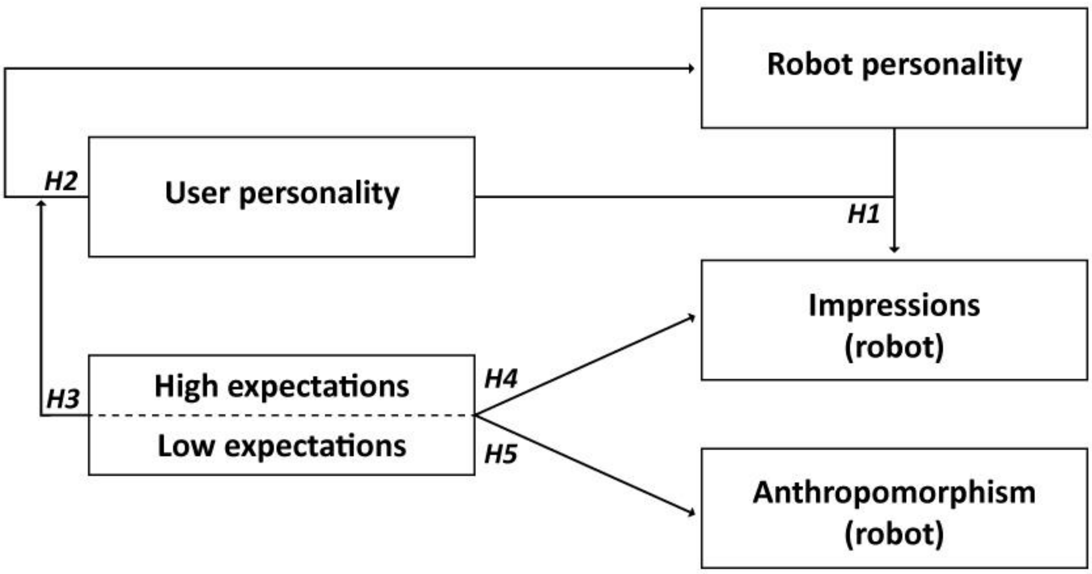
We also expect that expectation setting will influence the described phenomenon of anthropomorphism. We assume that humans with high expectations will anthropomorphise a social to a higher degree than humans with low expectations. Therefore, we have formulated the following hypothesis:

H5: Humans who have high expectations of a social robot will anthropomorphise this robot more than humans who have low expectations of a social robot.

Research Model

The five hypotheses presented in this study can be visualized into a research model. Figure 1 illustrates this research model and shows the relation between the central concepts within this study and indicates which relation each hypothesis seeks to clarify.

Figure I. Conceptual research model



Method

Design

In a between-participants experiment (high expectations vs. low expectations), we studied if expectation setting would influence the participants impressions of a social robot, if participants tend to assign their own personality traits to that of the social robot and if expectation setting would influence the extent to which we anthropomorphise a social robot.

Participants that were assigned to the “high expectations” condition received an instruction and information sheet about the social robot which described the robot as a living pet. Participants assigned to the “low expectations” condition received an instruction and information sheet about a social robot which described the robot as a kind of stuffed animal (Appendix A).

Experiment Manipulation

The experiment manipulation within this study consisted of user expectation setting, whereby the participants could have high or low expectations. This manipulation of the participants expectations of the social robot was done by an information sheet with text about the social robot and a picture of the robot. The experimenter emphasized that the participant had to read the instruction sheet carefully. In order to not only depend on the written briefing for expectation setting, the experimenter also emphasized the highlights from the briefing verbally. Some of the highlights which were formulated in the briefing and mentioned by the experimenter in the high expectation setting were: “New product”, ”completely autonomous”, ”lifelike pet”, “has learning capabilities” and “has multiple motors and many sensors”. In the low expectation setting these terms were like: “Outmoded robot”, ”like a stuffed animal”, ”limited engine capacity”, “no learning capabilities” and “limited number of sensors”. An example of the instruction sheets can be found in Appendix A. In order to manipulate expectations successfully, it was important that participants in both the low and high expectations condition received the same amount of stimuli. Therefore we used opposite arguments in the instruction sheets. We also investigated how different studies in the past manipulated expectations.

Paepcke and Takayama (2010) conducted an experiment in which they manipulated expectations. The experimenter within this study informed the participants verbally about the functions and specifications of the robot with terms belonging to either the high expectations condition or low expectations condition. They also used signs to emphasize the mentioned highlights. These signs could be compared with our information sheet, which the participants used during the interaction. Duffy (1986) conducted a study in which expectations were manipulated by using text. Expectations were manipulated by using 2 different texts with 40 short narratives, which included high or low expectations. In both studies the manipulation turned out to be successful. Therefore, we can conclude that expectation setting on the basis of text and speech seems to be a workable solution for manipulating expectations.

Materials

The social robot which has been used in this study is the zoomorphic robot Pleo. Pleo looks like a small dinosaur and has approximately the size of a cat. He can act autonomously, explores and reacts to the environment, interacts with humans and express emotion. The skin of Pleo consists of a rubber texture which covers a mechanical frame. Pleo runs on fourteen different motors which are placed in different segments of his body, through what he can shake his tail, bend his neck in different positions, control his mouth and eye-lid and walk slowly. Pleo has the capability of making different noises, which express his feelings. It also has many different sensors all over his body, including eight capacitive touch sensors, two infrared (IR) sensors and a small CMOS camera, which is mounted on the nose (Fernaues et al., 2010).

Pleo has a unique feature which makes it possible to install different programmed personalities and operating systems. In this study we have upgraded Pleo to LifeOS 2.0.1, which is an improved version of the Pleo software. With LifeOS 2.0.1 Pleo is more active, has more opportunities to grow and learn, has new tricks and sounds and all of the updates from LifeOS 1.1 are included. After LifeOS 2.0.1 was installed, Pleo has been raised to its last stadium, whereby he has a fixed personality.

Measures

For measuring the personality of the robot and the participant we have used the 40-item Mini-Marker Set that is derived from the Big-Five inventory set (Goldberg, 1992). Saucier (1994) developed this 40-item Mini-Marker Set and it can be used for measuring the different factors of personality, namely: Extraversion, Agreeableness, Conscientiousness, Emotional Stability and Openness. The items of this set where translated to a Dutch version by using the translations of Denissen et al. (2008). Each item could be rated from a number one to seven, whereby a one represented an extremely inaccurate personality trait and a seven an extremely accurate personality trait. We have used the items of Bartneck et al. (2009) in order to measure the degree to which the participants anthropomorphised Pleo. These items were placed on a 7-point semantic differential: Fake/Natural, Machinelike/Humanlike, Unconscious/Conscious, Artificial/Lifelike, and Moving rigidly/Moving elegantly.

The impressions that participants had of Pleo after interaction, were measured with the SD method proposed by Osgood et al. (1957). Impressions are ideas, feelings, or opinions about something or someone, especially one formed without conscious thought or on the basis of little evidence (Oxford Dictionaries, n.d.). The SD method, which has been used in the past for measuring impressions of robots (Nishimura et al. 2005), exists of twenty-eight adjective pairs which has been placed on a 7-

point semantic differential (Kanda et al. 2001). Factor analysis was performed on the SD method ratings for the 28 adjective pairs. It was observed that all 28 items correlated at least .45 with at least one other item, suggesting reasonable factorability (Appendix B). The Kaiser-Meyer-Olkin measure of sampling adequacy was .854, above the commonly recommended value of .6, and Bartlett's test of Sphericity was significant (X^2 (378) = 1604.93, $p < .001$). Given these overall indicators, factor analysis was deemed to be suitable. Based on the difference in eigenvalues, we adopted a solution that consists of 3 factors. Cumulative proportion of these factors was 57.4%. The retrieved factor matrix was rotated by a Varimax method (Appendix B). We named the 3 factors: Familiarity impressions, performance impressions and activity impressions. These factor names are derived from Kanda et al. (2001), which found four factors. They named them familiarity factor, enjoyment factor, performance factor and activity factor. In our factor analysis, we moved the items of the enjoyment factor to the familiarity factor, because factor analysis revealed better factor loadings when the familiarity and enjoyment items of Kanda et al. (2001) were combined as one factor. We kept the name familiarity, because we assume that enjoyment can be related to familiarity. Enjoyment can be related to familiarity, because research in the past have recognized that familiarity is an important predictor of enjoyment (Schubert, 2007). The familiarity factor represents terms like "Kind", "Accessible", "Exciting" and "Pleasant". The performance factor represents terms like "Intelligent", "Quick" and "Sharp". The activity factor represents terms like "Agitated", "Showy" and "Cheerful". The Cronbach's Alpha scores for the different impressions constructs were all above $\alpha = 0.75$ (Table *I*), which means that constructs had a good internal consistency.

We also measured the participants interest in technology. Therefore we used the Personal Innovativeness In Technology (PIIT) scale, which reflects the willingness of a person to try out new technologies (Agarwal & Prasad, 1998). The statements on the PIIT scale could be rated from a number one to seven, whereby a one represented an extremely inaccurate statement description of the participant and a seven an extremely accurate description. The different Cronbach's Alpha scores for the different constructs can be found in table I. As the Cronbach's alpha in this study were all much higher than 0.6, the constructs were therefore deemed to have an acceptable reliability (Sekaran, 1992).

Table I. Cronbach's Alpha of the constructs

Constructs	**Items**	**Original *a***	**Items removed**	**Resulting *a***
Personal Innovativeness In Technology	4	.83	0	.83
Extraversion (participant himself)	8	.83	0	.83
Agreeableness (participant himself)	8	.71	0	.71
Conscientiousness (participant himself)	8	.72	0	.72
Emotional Stability (participant himself)	8	.78	0	.78
Openness/Intellect (participant himself)	8	.69	0	.69
Anthropomorphism	5	.87	0	.87
Impressions (total)	28	.94	0	.94
Impressions (familiarity factor)	10	.90	0	.90
Impressions (performance factor)	11	.91	0	.91
Impressions (activity factor)	7	.77	0	.77
Extraversion (Pleo)	8	.66	0	.66
Agreeableness (Pleo)	8	.67	0	.67
Conscientiousness (Pleo)	8	.51	3	.69
Emotional Stability (Pleo)	8	.79	0	.79
Openness/Intellect (Pleo)	8	.81	0	.81

Procedure

Before participants could take part in this study, they were asked if they had ever interacted with the robot Pleo. Only the participants who never had interacted with Pleo before participated in this study. Firstly, the participants completed the 40-item Mini-Marker questionnaire for measuring their personality traits. After completing the questionnaire, they were briefed about the study, whereby the experimenter explained that the participants could interact freely with Pleo for approximately ten minutes. Freely interaction also meant that the participant had the possibility to stop the interaction within the ten minutes when they had sufficient impressions of the social robot Pleo. However they first had to read the information and instruction sheet (expectation setting) before they could interact with Pleo. The experimenter repeated some key instructions verbally after the participant had read the instruction sheet and he explained how the participant could interact with Pleo.

The participants had the possibility to freely interact with Pleo for approximately 10 minutes. The choice for this freely interaction instead of task based interaction was chosen in order to give the participants enough freedom for interacting with Pleo in their own way and to avoid disappointment. We assumed that task based interaction could lead to more disappointment when a participant failed to perform a specific task, therefore freely interaction seemed to be a better method.

When the participants finished interacting with Pleo they completed a second questionnaire where they had to assign a personality to the robot by filling in the 40-item Mini-Marker set for Pleo. This questionnaire also contained scales for measuring the impressions participants had of Pleo and scales for measuring the extent to which the participants anthropomorphised Pleo. The questionnaire ended with the question if the participants knew what the purpose of this study was. Once the participants completed the last questionnaire, they were briefed about the real purpose of this research.

Participants

A total of 86 respondents participated in this experiment. The characteristics of the participants are illustrated in table II. Participants were recruited within the University of Twente and by using the network of the researcher. The majority of the participants were obtained from the Behavioural Sciences department.

The personality traits of these participants could be characterized as follows: Extraversion (M = 4.92, SD = 0.95), Agreeableness (M = 4.72, SD = 0.75), Conscientiousness (M = 4.83, SD = 0.70), Emotional Stability (M = 4.42, SD = 0.91) and Openness (M = 4.84, SD = 0.72). The participants could only participate in this study if this was the first time that they would interact with the social robot Pleo, because we assume that earlier interactions with a Pleo could influence the expectation setting. However, none of the total participants was familiar with Pleo, therefore nobody was excluded from this study. Participants were also required to have a minimum age of 16 years old. The participants were not told of the purpose of the study until the end of the experiment. Some participants received so called credit points, which they need for their education by participating in different experiments.

Table II. Sample characteristics

Sample characteristics		**Frequency**	**Percent**
Gender	Male	41	47,7%
	Female	45	52,3%
Age	16 - 20 years	11	12,8%
	21 - 25 years	45	52,3%
	26 - 30 years	20	23,3%
	31 - 35 years	3	3,5%
	36 years and older	7	8,1%
Nationality	Dutch	82	95,3%
	German	4	4,7%
Education	Primary education	0	0,0%
	Secondary education	1	1,2%
	Intermediate Vocational Education	19	22,1%
	Higher Vocational Education	19	22,1%
	Academic education	47	54,7%

Results

Manipulation Check

By means of an open-ended question we tested if participants knew the purpose of this study and if they realized that they were manipulated. We can conclude that our manipulation of setting high or low expectations was successful, because no one mentioned that their expectations were manipulated. In fact, no one even mentioned that they were part of a specific research condition. However, some participants (N=8) described that this study measured personality traits and 2 participants described that the research was about comparing these personality traits. We did not exclude these participants from our study, because their awareness of the purpose of this study was to minimal. Therefore, we assumed that it would not influence the results.

Effects of a matching personality on impressions

We first calculated the differences between participants' personality traits and assigned personality traits to Pleo. There was a matching personality when the difference was between minus one and one. A One-Way analysis of variance was conducted to examine Hypothesis 1, which described that personality matching would influence the total impressions a participant had of the social robot Pleo. A Conscientiousness personality was the only personality trait that influenced the participants total impressions of the social robot Pleo in case of a matching personality. The participants with a matching conscientiousness personality had an average total impression for Pleo of 4.89 (SD = 0.88). Participants which had not a matching conscientiousness personality had an average total impression for Pleo of 4.22 (SD = 0.62). The effect of a matching conscientiousness personality on total impressions, therefore was significant, $F(1, 84) = 15.817, p < .001$. This means that participants with a matching conscientiousness personality had more positive total impressions of the social robot Pleo, then participants that had a non-matching conscientiousness personality. Therefore, Hypothesis 1 was partially supported.

Further analysis was performed to study the effects of personality attribution on the total impressions. We used dichotomizing to split the continuous variables into categorical variables that had two levels. For example: We categorized the subjects as either Introvert (1 through 3,5 on the scale) or Extravert (3,5 through 7 on the scale). Within this further analysis we used dichotomizing, because it makes an ANOVA analysis possible (Kline, 2009, p. 49), it simplifies the presentation of results and produces meaningful findings that are easily understandable to a wide audience (Farrington & Loeber, 2000). The results of a Two-Way analysis of variance indicated that some main effects existed. The main effect of an extraverted/introverted personality of Pleo was significant, $F(1, 82)= 4.386$, $p < .05$. This indicates that participants had significant more positive impressions of Pleo when it had an assigned extraverted personality (M = 4.70, SD = 0.84) in comparison with an assigned introverted personality (M = 4.21, SD = 0.76). The main effect of an agreeable personality of Pleo was marginally significant, $F(1, 82) = 3.662$, $p = .059$. This indicates that participants had significant more positive impressions of Pleo when they assigned high levels of agreeableness to Pleo ($M = 4.67$, $SD = 0.82$) in comparison with assigned low levels of agreeableness ($M = 3.86$, $SD = 0.82$). No further main effect were found.

There was no significant interaction between participants own personality and assigned personality to Pleo on the impressions that participants have of Pleo, Extraversion Pleo ($F(1, 82)= 1.444$, $p > .05$), Agreeableness ($F(1, 82)= 0.086$, $p > .05$), Conscientiousness ($F(1, 82)= 0.072$, $p > .05$), Emotional Stability ($F(1, 82)= 0.401$, $p > .05$) and Intellect/Openness ($F(1, 82)= 0.104$, $p > .05$).

Assigning own personality to a social robot

Pearson correlation coefficients were calculated to investigate the relationship between the assigned personality traits of the participant and Pleo. We wanted to examine Hypothesis 2, which described that participants tend to assign their own personality traits to a social robot. Only one significant positive correlation was revealed between subjects Agreeableness and the assigned Agreeableness to Pleo, $r(84) = .209$, $p < .05$. Linear regression analyses on this correlation was used to specify the nature of the relation between participants own agreeable personality and the assigned agreeable personality of Pleo. Participants own agreeable personality and assigned agreeable personality to Pleo had marginal significant positive regression weights ($\beta = 0.241$, $t = 1.96$, $p = 0.053$). This indicates that the more a participant assigned an agreeable personality to himself/herself on the first questionnaire, the more a participant was expected to assign an agreeable personality to Pleo at the second questionnaire. Therefore Hypothesis 2 was partially supported.

Effects of expectation setting on assigning personality to a social robot

Pearson correlation coefficients were calculated to investigate if expectation setting would influence the relationship between the assigned personality traits of the participant and Pleo. Hypothesis 3 described that participants with high expectations tend to assign their own personality traits to a social robot more than participants with low expectations. We found a marginal significant negative correlation for participants with low expectations between subjects intellect/openness personality traits and the assigned intellect/openness personality traits to Pleo, $r(84) = -.217$, $p = .081$. This would indicate that the more a participant assigned intellect/openness personality traits to himself/herself on the first questionnaire, the less a participant assigned intellect/openness personality traits to Pleo at the second questionnaire within the low expectations condition. However, linear regression analyses on this correlation was non-significant ($\beta = -0.339$, $t = -1.422$, $p > 0.05$). This means that participants own intellect/openness personality traits did not explain the level of an assigned intellect/openness personality to Pleo.

We also found a significant positive correlation and regression for participants which had low expectations. A significant positive correlation was revealed within the low expectations condition between subject agreeableness personality traits and the assigned agreeableness personality traits to Pleo, $r(84) = -.447$, $p < .01$. This indicates that the more a participant assigned agreeableness personality traits to himself/herself on the first questionnaire, the more a participant assigned agreeableness personality traits to Pleo at the second questionnaire within the low expectations condition. Further analyses confirmed this statement, because regression analyses was significant ($\beta = 0.545$, $t = 3.200$, $p < 0.01$). No further significant correlations were found between participant Extraversion, Conscientiousness and Emotional Stability personality traits and Pleo Extraversion, Conscientiousness and Emotional Stability personality traits within the two expectation conditions. Because we only found effects of attributive projection in the low expectations setting, we reject hypothesis 3.

Effects of expectation setting on impressions of a social robot

Hypothesis 4 stated that expectation setting would influence a participants' impression of the social robot. A One-Way analysis of variance was conducted to examine if expectation setting would influence the total impressions that a participant had of Pleo. Results from a One-Way analysis of variance indicate that expectation setting affected the impressions a participant had of the social robot after he/she interacted with Pleo. The participants in the high expectations group had an average total impression for Pleo of 4.96 ($SD = 0.88$). The participants in the low expectations group had an average total impression for Pleo of 4.24 ($SD = 0.62$). The effect of having low or high expectations on total impressions was significant, $F(1, 84) = 19.55$, $p < .001$. Therefore Hypothesis 4 was supported, which means that participants within the high expectation condition, after interaction, had more positive total impressions of the social robot Pleo in comparison with participants in the low expectations condition.

Further analysis of the different factors within the total impressions, which are familiarity, performance and activity impressions, also indicate that expectation setting affected the impressions someone had of Pleo. Participants within the high expectations condition had marginal significantly more positive familiarity impressions of Pleo ($M = 5.62$, $SD = 0.93$) than participants within the low expectations condition ($M = 5.27$, $SD = 0.79$), $F(1, 84) = 3.485$, $p = .065$. Participants within the high expectations condition had significant more positive performance impressions of Pleo ($M = 4.37$, $SD = 1.06$) than participants within the low expectations condition ($M = 3.36$, $SD = 0.87$), $F(1, 84) = 23.331$, $p < .001$. Participants within the high expectations condition had significant more positive activity impressions of Pleo ($M = 4.95$, $SD = 1.03$) than participants within the low expectations condition ($M = 4.10$, $SD = 0.81$), $F(1, 84) = 18.051$, $p < .001$. We can confirm that expectation setting had the greatest effect on performance and activity impressions. To conclude, hypothesis 4 was confirmed.

We also performed a Two-Way analysis of variance to examine the effect of gender and expectations on the total impressions of the social robot Pleo. The main effect of gender was non-significant, $F(1, 82) = .208$, $p > .05$. This indicates that gender did not affect the total impressions participants had of Pleo. However, there was a marginal significant interaction effect between gender and expectations, $F(1, 82) = 3.17$, $p < .1$. This interaction effect indicates that the gender effect was greater in the low expectations condition than in the high expectations condition. The effect of gender on total impressions in the low expectations condition was marginal significant, $F(1, 41) = 3.965$, $p = .053$. This means that women, in the low expectations condition, had more positive total impressions ($M = 4.41$, $SD = 0.66$) of Pleo than men ($M = 4.05$, $SD = 0.53$).

Effects of expectation setting on anthropomorphising a social robot

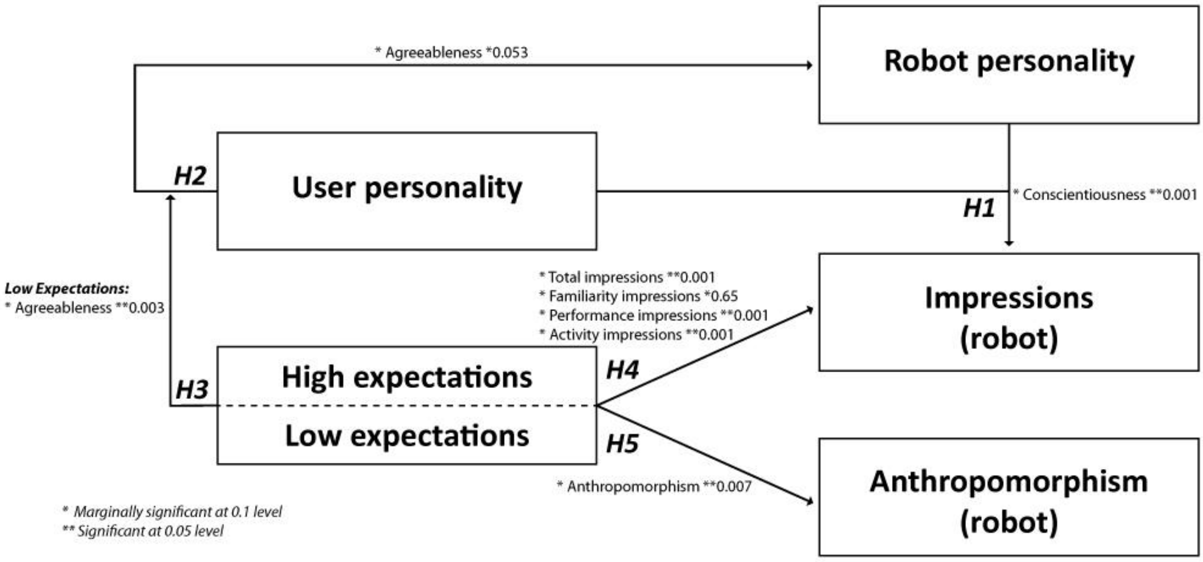
A One-Way analysis of variance was conducted to examine Hypothesis 5, which described that expectation setting would influence the degree to which the participants anthropomorphised the social robot Pleo. The results indicated that expectation setting affected the degree to which the participants anthropomorphised the social robot Pleo. The participants in the high expectations condition anthropomorphised the social robot Pleo more ($M = 3.68$, $SD = 1.48$) than participants in the low expectations condition ($M = 2.95$, $SD = 0.88$). The effect of having high or low expectations on anthropomorphising was significant, $F(1, 84) = 7.599$, $p < .01$. This means that Hypothesis 5 was supported.

Further analysis was performed to study the effects of Personal Interest In Technology on anthropomorphising a social robot. We used dichotomizing to split the continuous variable into a categorical variable that had two levels, namely: Low interest in technology (1 through 3,5 on the scale) and High interest in technology (3,5 through 7 on the scale). Then we conducted a Two-Way analysis of variance to examine if participants Personal Interest In Technology (PIIT) and expectations influenced the degree to which participants anthropomorphised Pleo. The main effect of PIIT was not significant $F(1, 82) = 2.900$, $p = .317$. This means that participants personal interest in technology did not influence the degree to which they anthropomorphised the social robot. However, we found that in the low expectations condition, participants with low interest for technology ($M = 3.24$, $SD = 0.97$) significant anthropomorphised Pleo more ($F(1, 41) = 6.838$, $p < .05$) than participants with high interest for technology ($M = 2.68$, $SD = 0.71$). There was no significant interaction or main effect for gender.

Results processed in research model

The outcomes which can be directly linked to the five hypotheses are illustrated in figure II. In this figure we have listed all the significance levels of the different results and placed them at the corresponding hypothesis.

Figure II. research model with significant levels



Discussion and Conclusion

In this study we investigated if expectations of a social robot would influence the impressions humans had of a social robot, the degree to which humans anthropomorphised a social robot and to what extent humans would assign their own personality traits to that of the social robot. We also studied the influence of matching personality traits, between human and robot, on the impressions that this human had of the social robot. In summary, we found that humans had more positive impressions of a social robot when there was a matching conscientiousness personality. Humans, especially within the low expectations condition, which describe themselves as being agreeable tend to assign this personality trait to a social robot. Humans which had high expectations had more positive impressions of a social robot and anthropomorphised a social robot to a higher degree in comparison with humans which had low expectations.

Hypothesis 1 stated that humans would have more positive impressions of a social robot when the robots' personality matched with that of the human. The results from our study indicate that a matching personality between user and robot influenced the total impressions one had of the social robot. Although, this was only the case for the conscientiousness personality traits. Humans had more positive impressions of a social robot when he/she assigned a matching conscientiousness personality to a social robot. This finding attributes to and supports the similarity attraction hypothesis for a conscientiousness personality. This result is consistent with research in past which studied the similarity-attraction hypothesis in the field of human-robot interaction (Bernier & Scassellati, 2010; Park et al., 2012). However, there was no evidence that the similarity-attraction hypothesis held true for the Extraversion, Agreeableness, Emotional stability or Intellect/Openness personality traits. Nevertheless, this did not mean that the complementary principle could be related to these personality traits and the impressions one had of a social robot, because humans with a matching or non-matching extravert, agreeable, emotional stable or intellect/openness personality had the same impressions of the social robot. These results are in contrast with research on the complementarity principle in the field of human-robot interaction (Lee et al., 2006). We also found evidence that an assigned extravert or agreeable personality to a robot influenced the total impressions one had of the robot. This implies that robot users had more positive impressions of a social robot when they attributed an extravert or agreeable personality to the social robot. Future studies can use this knowledge when they need information about what kind of personality a robot should manifest, for example: A study in which researchers need to program a robot personality that must evoke positive impressions among humans. So to conclude, the findings partially support hypothesis 1 and therefore this hypothesis was partially confirmed.

Hypothesis 2 stated that humans tend to assign their own specific personality traits to a social robot. This hypothesis was derived from the assumed similarity theory, which refers to the belief that others are similar to the self (Cronbach, 1955). The attribution of own personality traits to others is also known as attributive projection (Holmes, 1978). Our results indicate that hypothesis 2 was partially supported. The more a participant described himself/herself as having an agreeable personality, the more he/she described the robot as being an agreeable robot. This finding expands the theory of assumed similarity and attributive projection for the agreeableness personality trait, within the field of human-robot interaction. However, there was no relationship between extravert, conscientiousness, emotional stable and intellect/open personality traits of a robot user and the attribution of these personality traits to a social robot. Srivastava et al. (2010) also supported this hypothesis, by stating that the more a person described himself of having a certain personality trait, the more they will dedicate this personality traits to others. Our results are partially consistent with that of Woods et al. (2005), which investigated this phenomenon within the field of human-robot interaction. They found that younger subjects tend to ascribe their own personality traits to the social robot. We can compare this study of Woods et al. (2005), because our study mainly existed of students. So to conclude, the findings partially support hypothesis 2 and therefore this hypothesis was partially confirmed.

Hypothesis 3 investigated if expectation setting had an effect on the attribution of specific personality traits to a social robot. This hypothesis had much consistency with hypothesis 2, but now we added the possible influence of expectation setting to this hypothesis. We stated that humans with high expectations of a social robot, compared to humans with low expectations, were more likely to assign their own personality traits to a social robot. This hypothesis was derived from our suggestion that humans with high expectations of a zoomorphic robot are more likely to see the robot as a living creature and therefore they will be more inclined to project their own personality traits to a social robot. Halpern and Goldschmitt (1976) stated that attributive projection occurs, because this enables humans to see others as more familiar and less threatening. We expanded this assumption by stating that humans with high expectations will have more motivations to see their interaction partner as familiar and less threatening. Our findings suggest that attributive projection occurred, but this was only the case in the low expectations setting. There was no evidence that attributive projection occurred in the high expectation setting. Attributive projection occurred in the low expectations condition for personality traits related to agreeableness. This means that agreeable humans with low expectations of a social robot were more likely to assign an agreeable personality to the social robot. We also found contrary results within the low expectations setting for personality traits related to intellect/openness. Humans with low expectations and a intellect/openness personality were not likely to attribute this personality trait to a social robot. They tended to assign an ignorant/closed personality to the social robot. However, regression analyses stated that there was not a positive linear relationship. So to conclude, the findings did not support hypothesis 3 and therefore this hypothesis was not confirmed.

Hypothesis 4 suggested that humans with high expectations of a social robot would have more positive impressions of that robot than humans with low expectations. Theories about the Self-Fulfilling Prophecy and the Confirmation Bias provided the basic principles within this hypothesis. Both theories can be related to expectation setting, because they explain how specific expectations influence our behaviour towards objects, robots and other people. These theories predict that when humans have high expectations, they will seek evidence to live up to these expectations. A more recent study suggested that expectations for a social robot should be set low, because they found evidence that humans with low expectations had more positive perceptions of a robot in comparison with humans which beforehand had high expectations (Paepcke & Takayama, 2010). However, we argue that their sample size (N = 24) is too small for drawing this conclusion. Therefore knowledge concerning expectations setting and human-robot interaction is scarce. Our study differed from that of Paepcke and Takayama (2010), in which we measured impressions and they measured perceptions. When participants have a small amount of human-robot interaction time, then it is a better solution to measure impressions instead of perceptions, because impressions are formed on the basis of less evidence than perceptions (Oxford Dictionaries, n.d.). So it was possible that we would find opposite results. Nevertheless, our results are consistent with the Self-Fulfilling Prophecy and Confirmation Bias. We found evidence that humans with high expectations of a social robot had more positive impressions of this robot in comparison with humans which had low expectations. The total impressions a participant had of the social robot could be divided in three categories, namely: Familiarity impressions, Performance impressions and Activity impressions. Humans with high expectations had more positive familiarity, performance and activity impressions of the social robot, than humans with low expectations. Expectations setting had the greatest effect on performance and activity impressions. These findings could be seen as an important theoretical implication within the field of human-robot interaction, whereby future human-robot research must be aware of the effect of expectations on the results. Further analysis with regard to gender and total impressions indicated that there was no difference between women and men in the high expectations group. Women in the low expectations group had more positive impressions of the social robot than men. So to conclude, all these findings support hypothesis 4 and therefore this hypothesis was confirmed.

We also found evidence which supported hypothesis 5, which stated that humans with high expectations of a social robot would anthropomorphise the robot more than humans with low expectations. This hypothesis is derived from our argumentation that humans with high expectations are more likely to judge a social robot as a living creature than humans with low expectations. This argumentation turned out to be true, so humans with high expectations will anthropomorphise a social robot to a higher degree than humans with low expectations. This result can be explained by the Three-Factor-Theory of Anthropomorphism (Epley et al., 2007). In this theory the authors describe that Effectance Motivation influences the extent to which humans anthropomorphise. Based on our results we assume that humans with high expectations have more motivations than humans with low expectations to interact effectively with a social robot and therefore they will attribute more human/animal characteristics to a social robot. We also measured the Personal Interest In Technology of the participants, because it would be interesting to find differences between participants with high interest in technology and participants with low interest in technology on the degree to which they anthropomorphised the social robot. It turned out that participants with low expectations and low interest in technology anthropomorphised the social robot to a higher degree than people with low expectations and high interest in technology. So to conclude, all these findings support hypothesis 5 and therefore this hypothesis was confirmed.

Beside contributing knowledge to field of human-robot interaction, we also developed a Dutch semantic differential for measuring impressions. The English semantic differential with 28 adjective pairs was translated to Dutch (Kanda et al. 2001). Kanda et al. (2001) found four factors within this semantic differential, namely: Familiarity factor, Enjoyment factor, Performance factor and Activity factor. Our Dutch semantic differential for measuring impressions, did measure the robot user impressions adequately and factor analysis performed on these 28 adjective pairs revealed three factors, namely: Familiarity impressions, Performance impressions and Activity impressions. So in our study we moved the enjoyment factor to the familiarity factor, because we assume that enjoyment can be related to familiarity. Research in the past have recognized that familiarity is an important predictor of enjoyment (Schubert, 2007). A possible explanation for the fact that we found three factors instead of four is that we used a robot that was developed for the consumer market. Therefore, we assume that it is possible that enjoyment and familiarity impressions are more equal when humans are confronted with a robot that is developed to interact with humans as if it is a real living creature. Kanda et al. (2001) used a robot that was only operational in a laboratory setting. The Cronbach's Alpha of the familiarity, performance and activity factors indicated that the items within these factors measured the same concept. Therefore future research that will measure impressions within the Netherlands can make use of this scale.

Practical implications

Besides contributing scientific knowledge to the field of human-robot interaction, the gathered results can also provide insights on how a social robot should be marketed. This study provides social robot manufactures and more specific zoomorphic robot manufactures knowledge about how they should market and position the social robot. We suggest that these manufactures should market and position their social robots in such a way that it will shape high expectations, because we found evidence that humans with high expectations are more likely to have positive impressions of a social robot than humans with low expectations. Social robots are relatively new to humans and they could see social robots as “the next big thing”, therefore it is important that these robots will be positioned as advanced technological products, because than humans will be more inclined to judge the robot as capable. Another practical implication is that shaping high expectations will increase the degree to which robot users anthropomorphise the social robot. Therefore, we also suggest that zoomorphic robot manufactures should market their products as if these are real living creatures which can be compared with pets, because anthropomorphism helps humans to fulfil their need of being social (Breazeal, 2002).

This study helps robot programmers and personality developers to determine what kind of personality traits a robot should manifest. In case of a zoomorphic robot, the robot should manifest high levels of extraversion and/or agreeableness.

Limitations and future work

There are several limitations within this study that could be improved in future studies. First, we only used one robot in this study, therefore it is difficult to generalize the findings to other different social robots. Besides, the used robot was introduced to consumers in 2008 and is a somewhat outmoded social robot. Future research on expectation setting and personality attribution for social robots should include not only a zoomorphic social robot but also a humanoid social robot. Such a study can investigate if human-robot interaction differs between a zoomorphic and a humanoid social robot in the context of expectation setting and the influence it has on impressions and personality attribution.

The subjects within this study mainly existed of students and therefore we have a good sample for predicting how expectation setting influences the younger generation. However, it is not sure if these findings also can be generalized to the older people. Future research should compare these two groups and investigate if they differ from each other concerning the influence of expectation setting on impressions of a social robot, anthropomorphising a social robot and assigning personality to a social robot.

Although, our manipulation check suggested that our expectation setting was successful, we could not prove statistically that we have manipulated expectations successfully. There is still a chance that not every subject had the same high or low expectations for the social robot. In order to set approximately the same high expectations for every subject within this condition and the same low expectations for subjects within this condition, future experiments should consider if they want to use more comprehensive techniques for expectation setting, such as television commercials, advertisements and/or user reviews. Future studies should also include measures about the expectations a participant had of the social robot, because then it is possible to make an even better distinction between high and low expectations. Within this study we defined expectations and impressions by using an accepted online dictionary, because we could not find an acceptable scientific definition of these terms. We suggest that future studies should determine a clear and comprehensive definition of expectations and impressions.

Conclusion

To conclude, results from our experiment indicate that the similarity-attraction hypothesis was only supported for a conscientiousness personality. This means that a similar or complementary conscientiousness robot personality influenced the impressions of a social robot positively. We also found that humans had more positive impressions of a social robot, when this robot had an assigned extravert or agreeable personality. The theory of assumed similarity and projective attribution between user and the robot, was partially supported. We found some evidence that assumed similarity or attributive projection occurs when humans describe themselves as having high levels of agreeableness. Further analysis indicated that expectations setting also influenced the occurrence of assumed similarity or attributive projection. Humans with low expectations and an agreeable personality, are more likely to assign an agreeable personality to a social robot. On the other hand, humans with low expectations and an intellect/openness personality, are less likely to attribute this personality trait to a social robot and they tend to describe the social robot as unintelligent/closed.

There was no evidence that assumed similarity or attributive projection occurred in the high expectations setting.

Finally, the suggestion that humans with high expectations would have more positive impressions of a social robot and would anthropomorphise this social robot to a higher degree in comparison with humans which have low expectations, turned out to be true. This means that future robot introductions should emphasize all the qualities a robot has, because then humans are more likely to have positive impressions of social robot and they will attribute more humanlike properties, characteristics and mental states to the social robot.

Literature

Agarwal, R., & Prasad, J. (1998). A conceptual and operational definition of personal innovativeness in the domain of information technology. *Information systems research, 9*(2), 204-215.

Bartlett, B., Estivill-Castro, V., & Seymon, S. (2004). Dogs or robots: why do children see them as robotic pets rather than canine machines? *Fifth Australasian User Interface Conference.* Dunedin. Conferences in Research and Practice in Information Technology, *28*, 7-14.

Bartneck, C., & Forlizzi, J. (2004). A design-centered framework for social human-robot interaction. *Proceedings of the Ro-man,* 591-594. DOI: 10.1109/ROMAN.2004.1374827.

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics, 1*(1), 71-81.

Beck, A.M., Edwards, N.E., Kahn, P.H., & Friedman, B. (2004). Robotic pets as perceived companions for older adults. *Tenth International Conference on Human-Animal Interactions*.

Bernier, E.P., & Scassellati, B. (2010). The similarity-attraction effect in human-robot interaction. *Ninth International conference on development and learning,* 286-290.

Breazeal, C. (2002). *Designing Sociable Robots.* The MIT Press.

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

Burgoon, J.K., & Hale, J.L. (1988). Nonverbal expectancy violations: Model elaboration and application to immediacy behaviours. *Communications Monographs, 55*(1), 58-79.

Byrne, D., Clore, G.L., & Smeaton, G. (1986). The attraction hypothesis: Do similar attitudes affect anything? *Journal of Personality and Social Psychology, 51,* 1167– 1170.

Costa, P.T., & McCrae, R.R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13,* 653-665.

Cronbach, L.J. (1955). Processes affecting scores on "understanding of others" and "assumed similarity". *Psychological bulletin, 52* (3), 177-193.

Dautenhahn, K. (2007). Methodology and themes of human-robot interaction: a growing research field. *International Journal of Advanced Robotic Systems, 4*(1), 103-108.

Denissen, J.J., Geenen, R., van Aken, M.A., Gosling, S.D., & Potter, J. (2008). Development and validation of a Dutch translation of the Big Five Inventory (BFI). *Journal of Personality Assessment, 90*(2), 152-157.

Dryer, D.C. (1999). Getting personal with computers: how to design personalities for agents. *Applied Artificial Intelligence, 13*(3), 273-295.

Duffy, S.A. (1986). Role of expectations in sentence integration. *Journal of experimental Psychology: Learning, Memory, and Cognition, 12*, 208-219.

Epley, N., Waytz, A., & Cacioppo, J.T. (2007). On seeing human: a three-factor theory of anthropomorphism*. Psychological review, 114*(4), 864-886.

Expectation. (2013). In *Merriam-Webster.com*. Retrieved February 7, 2013, from http://www.merriam-webster.com/dictionary/expectation.

Farrington, D.P., & Loeber, R. (2000). Some benefits of dichotomization in psychiatric and criminological research*. Criminal Behaviour and Mental Health,10*(2), 100-122.

Fernaeus, Y., Håkansson, M., Jacobsson, M., & Ljungblad, S. (2010). How do you play with a robotic toy animal?: A long-term study of Pleo. *Proceedings of the 9th international Conference on interaction Design and Children,* 39-48.

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

Friedman, B., Kahn, P.H., & Hagman, J. (2003). Hardware companions?: What online AIBO discussion forums reveal about the human-robotic relationship. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 273-280.

Goetz, J., & Kiesler, S., (2002). Cooperation with a robotic assistant. *Proceedings of the CHI'02 Conference on Human Factors in Computing Systems,* 578-579.

Goldberg, L.R. (1992). The development of markers for the Big-Five factor structure. *Psychological assessment, 4*(1), 26-42.

Halpern, J., & Goldschmitt, M. (1976). Attributive projection: Test of defensive hypotheses. *Perceptual and Motor Skills*, *42*(3), 707-711.

Holmes, D.S. (1978). Projection as a defense mechanism. *Psychological Bulletin, 85,* 677-688.

Hwang, J., Park, T., & Hwang, W. (2012). The effects of overall robot shape on the emotions invoked in users and the perceived personalities of robot. *Applied Ergonomics, 44*(3), 459-471.

Impressions. (2013). In *Merriam-Webster.com*. Retrieved February 7, 2013, from http://www.merriam-webster.com/dictionary/impressions.

Isbister, K., & Nass, C. (2000). Consistency of personality in interactive characters: Verbal cues, non-verbal cues, and user characteristics*. International Journal of Human-Computer Interaction, 53*(1), 251-267.

Kanda, T., Ishiguro, H., Ishida, T. (2001). Psychological analysis on human-robot interaction. *Proceedings of the 2001 IEEE International Conference on Robotics and Automation*, 4166-4173.

Kerepesi, A., Kubinyi, E., Jonsson, G.K., Magnusson, M.S., & Miklosi, A. (2006). Behavioral comparison of human-animal (dog) and human-robot (AIBO) interactions. *Behavioural Processes, 73*, 92-99.

Kline, R.B. (2009). *Becoming a behavioral science researcher: A guide to producing research that matters.* The Guilford Press.

Lee, K., Ashton, M.C., Pozzebon, J.A., Visser, B.A., Bourdage, J.S., & Ogunfowora, B. (2009). Similarity and assumed similarity in personality reports of well-acquainted persons. *Journal of Personality and Social Psychology, 96*(2), 460-472.

Lee, K.M., Park, N., & Song, H. (2005). Can a robot be perceived as a developing creature? *Human Communication Research, 31*(4), 538-563.

Lee, K.M., Peng, W., Jin, S., & Yan, C. (2006). Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human–robot interaction. *Journal of Communication, 56,* 754–772.

McCrae, R.R., & John, O.P. (1992). An introduction to the Five Factor Model and its applications. *Journal of Personality, 60,* 175-215.

Merton, R. K. (1948). The self-fulfilling prophecy. *Antioch Review,* 8, 193-210.

Mori, M. (1970). The uncanny valley. *Energy, 7*(4), 33–35.

Nakajima, H., Brave, S., Nass, C., Yamada, R., Morishima, Y., & Kawaji, S. (2003). The functionality of human-machine collaboration systems Mind model and social behavior. *Proceedings of the IEEE Conference on Systems, Man, and Cybernetics,* 2381-2387.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers*. Journal of Social Issues, 56*(1), 81–103.

Nass, C., Moon, Y., Fogg, B.J., Reeves, B., & Dryer, C. (1995). Can computer personalities be human personalities?. *Conference companion on Human factors in computing systems*, 228-229.

Nickerson, R.S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175.

Nishimura, Y., Kushida, K., Dohi, H., Ishizuka, M., Takeuchi, J., & Tsujino, H. (2005). Development and psychological evaluation of multimodal presentation markup language for humanoid robots. *Proceedings of 2005 5th IEEE-RAS International Conference on Humanoid Robots*, 393-398.

Osgood, C.E., Suci, G.J. & Tannenbaum, P.H. (1957). *The measurement of meaning.* Urbana: University of Illinois Press.

Paepcke, S., & Takayama, L. (2010). Judging a bot by its cover: an experiment on expectation setting for personal robots. *Human-Robot Interaction,* 45-52.

Park, E., Jin, D., & Pobil, A.P. (2012). The Law of Attraction in Human-Robot Interaction. *International Journal of Advanced Robotic Systems, 9*(35).

Rabin, M., & Schrag, J.L. (1999). First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics, 114*(1), 37-82.

Saucier, G. (1994). Mini-markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of personality assessment, 63*(3), 506-516.

Scheeff, M., Pinto, J., Rahardja, k., Snibbe, S., & Tow, R. (2002). Experiences with Sparky, a Social Robot. *Socially Intelligent Agents.* Springer US, 173-180.

Schubert, E. (2007). The influence of emotion, locus of emotion and familiarity upon preference in music. *Psychology of Music, 35*(3), 499-515.

Sekaran, U. (1992). *Research Methods for Business, 2nd ed.* New York: Wiley.

Srivastava, S., Guglielmo, S. & Beer, J.S. (2010). Perceiving others' personalities: examining the dimensionality, assumed similarity to the self, and stability of perceiver effects. *Journal of Personality and Social Psychology, 98,* 520-534.

Syrdal, D.S., Dautenhahn, K., Woods, S., Walters, M., & Koay, K.L. (2007). Looking good? Appearance preferences and robot personality inferences at zero acquaintance. *Multidisciplinary collaboration for socially assistive robotics: papers from the AAAI spring symposium,* 86-92.

Tapus, A., Ţăpuş, C., & Matarić, M.J. (2008). User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics, 1*(2), 169-183.

Woods, S. (2006). Exploring the design space of robots: Children's perspectives. *Interacting with Computers, 18*(6), 1390-1418.

Woods, S., Dautenhahn, K., Kaouri, C., Boekhorst, R., & Koay, K.L. (2005). Is this robot like me? Links between human and robot personality traits. *Proceedings of 2005 5th IEEE-RAS International Conference on Humanoid Robots,* 375-380.

Appendix A: Manipulating expectations

Low expectations:

INFORMATIE & INSTRUCTIE PAGINA

PLEO

De robot waarmee u straks kennis zult maken is al een aantal jaar verkrijgbaar. Zijn naam is Pleo en hij toont vele vergelijkingen met een knuffeldier die kan bewegen. Pleo kan niet zelfstandig handelen en hij reageert gelimiteerd op uw handelingen, dit betekent dat hij enkele taken kan uitvoeren en niet leert van de interactie die u met hem heeft. U kunt Pleo geluiden laten maken, u kunt hem verzorgen en hij kan zich beperkt voortbewegen. Pleo heeft namelijk enkele mogelijkheden om te communiceren (knuffeldier geluiden en lichaamstaal), voelen en horen.

De manier waarop Pleo zich kan bewegen is beperkt en houterig, omdat de inwendige motoren niet genoeg capaciteit bieden. Hij heeft een tweetal speakers, waarmee hij een aantal standaard geïnstalleerde geluiden kan produceren, zoals een knuffeldier dat kan. Daarnaast heeft Pleo op een aantal vaste plaatsen een sensor zitten, waardoor hij sommige aanrakingen zal voelen.

Kortom, Pleo is een wat oudere sociale robot met een verouderde technologie, maar biedt wel een aantal leuke mogelijkheden tot interactie.

In onderstaande afbeelding en uitleg kunt u zien hoe u met Pleo kan interacteren:

TIPS & TRICKS (Natuurlijk kun je zelf ook experimenteren met Pleo)

1. Aai Pleo op verschillende plaatsen en zie hoe hij reageert.
2. Laat Pleo wat eten door het blaadje voor zijn neus te houden.
3. Ga naast Pleo staan en maak op 10 tot 40 cm geluid en zie hoe hij reageert.
5. Aai Pleo een tijdje onder zijn kin en zie hoe vrolijk hij wordt.
6. U kunt natuurlijk ook gewoon kijken wat Pleo doet als u hem met rust laat.

(DOE A.U.B. NIET TE WILD MET PLEO)

High expectations:

INFORMATIE & INSTRUCTIE PAGINA

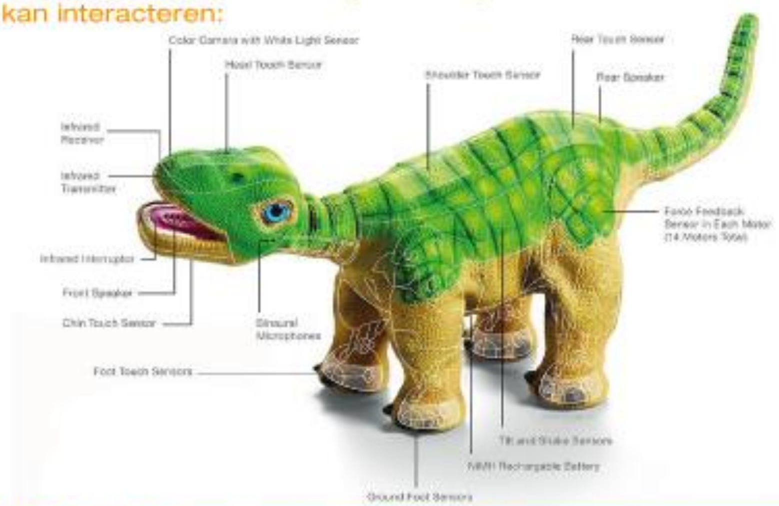
PLEO

De robot waarmee u straks kennis zult maken is nog maar net op de markt gebracht. Zijn naam is Pleo en hij toont vele vergelijkingen met een levensecht huisdier. Pleo kan zelfstandig handelen en hij reageert ongelimiteerd op uw handelingen, dit betekent dat hij vele taken kan uitvoeren en leert van de interactie die u met hem heeft. U kunt met Pleo communiceren, u kunt hem opvoeden en hij kan zich onbeperkt voortbewegen. Pleo heeft namelijk vele mogelijkheden om te communiceren (huisdier geluiden en lichaamstaal), voelen en horen.

De manier waarop Pleo zich kan bewegen is onbeperkt en levensecht, omdat de inwendige motoren genoeg capaciteit bieden. Hij heeft meerdere speakers, waarmee hij vrijuit verbaal met u kan communiceren, zoals een huisdier dat doet. Daarnaast heeft Pleo over zijn gehele lichaam sensoren zitten, waardoor hij alle aanrakingen van u zal voelen.

Kortom, Pleo is een moderne sociale robot met de nieuwste technologie, welke zeer realistische mogelijkheden biedt voor interactie.

In onderstaande afbeelding en uitleg kunt u zien hoe u met Pleo kan interacteren:



TIPS & TRICKS (Natuurlijk kun je zelf ook experimenteren met Pleo)

1. Aai Pleo op verschillende plaatsen en zie hoe hij reageert.
2. Laat Pleo wat eten door het blaadje voor zijn neus te houden.
3. Ga naast Pleo staan en maak op 10 tot 40 cm geluid en zie hoe hij reageert.
5. Aai Pleo een tijdje onder zijn kin en zie hoe vrolijk hij wordt.
6. U kunt natuurlijk ook gewoon kijken wat Pleo doet als u hem met rust laat.

(DOE A.U.B. NIET TE WILD MET PLEO)

Appendix B: factor analysis impressions

KMO and Bartlett's Test:

Kaiser-Meyer-Olkin Measure	of Sampling Adequacy	0,854
Bartlett's Test of Sphericity	Approx. Chi-Square	1604,934
	df	378
	Sig.	0

Factor matrix (Varimax rotated):

Items	Factor 1	Factor 2	Factor 3
Impressie wreed vs zachtaardig (Recode)	**0,80**	0,04	0,06
Impressie ongunstig vs gunstig (Recode)	**0,61**	0,30	0,22
Impressie onduidelijk vs duidelijk (Recode)	0,47	0,36	0,16
Impressie gevaarlijk vs veilig (Recode)	**0,77**	-0,06	0,15
Impressie kil vs hartelijk (Recode)	**0,83**	0,10	0,07
Impressie lelijk vs mooi (Recode)	0,44	**0,54**	0,15
Impressie gereserveerd vs spontaan (Recode)	**0,52**	0,47	0,24
Impressie onvriendelijk vs vriendelijk (Recode)	**0,78**	0,18	0,13
Impressie ontoegankelijk vs toegankelijk (Recode)	**0,66**	0,42	-0,17
Impressie kwaadgezind vs goedgezind (Recode)	**0,71**	0,17	0,23
Impressie egoistischvs onzelfzuchtig (Recode)	**0,52**	0,36	-0,18
Impressie mechanisch vs menselijk (Recode)	0,15	**0,71**	0,19
Impressie onvolledig vs volledig (Recode)	0,23	**0,75**	0,24
Impressie gezapig vs opwindend (Recode)	-0,04	**0,59**	0,40
Impressie onaangenaam vs aangenaam (Recode)	0,51	**0,54**	0,23
Impressie onaardig vs aardig (Recode)	**0,73**	0,21	0,20
Impressie saai vs interessant (Recode)	0,30	**0,76**	0,19
Impressie slecht vs goed (Recode)	0,45	0,44	0,47
Impressie eenvoudig vs complex (Recode)	0,09	**0,68**	-0,03
Impressie onintelligenvs intelligent (Recode)	0,25	**0,75**	0,19
Impressie langzaam vs snel (Recode)	0,02	**0,67**	0,38
Impressie kalm vs opgewonden (Recode)	-0,04	0,13	**0,76**
Impressie passief vs actief (Recode)	0,20	0,44	**0,55**
Impressie laf vs dapper (Recode)	0,40	0,35	0,37
Impressie terughoudend vs opvallend (Recode)	0,21	0,24	**0,57**
Impressie somber vs opgewekt (Recode)	0,50	0,08	**0,70**
Impressie stompziniigvs scherpzinnig (Recode)	0,44	**0,71**	0,09
Impressie traag vs vlug (Recode)	0,02	**0,48**	0,43

Rotation sums of eigenvalues and explained variance:

Factor	Eigenvalue	% of Variance	Cumulative %
Factor 1	6,702	23,935	23,935
Factor 2	6,234	22,263	46,198
Factor 3	3,137	11,203	57,401