ORIGINAL PAPER

Does the Design of a Robot Influence Its Animacy and Perceived Intelligence?

Christoph Bartneck · Takayuki Kanda · Omar Mubin · Abdullah Al Mahmud

Abstract Robots exhibit life-like behavior by performing intelligent actions. To enhance human-robot interaction it is necessary to investigate and understand how end-users perceive such animate behavior. In this paper, we report an experiment to investigate how people perceived different designs of robot embodiments in terms of animacy and intelligence. iCat and Robovie II were used as the two embodiments in this experiment. We conducted a between-subject experiment where robot type was the independent variable, and perceived animacy and intelligence of the robot were the dependent variables. Our findings suggest that a robot's perceived intelligence is significantly correlated with animacy. The correlation between the intelligence and the animacy of a robot was observed to be stronger in the case of the iCat embodiment. Our results also indicate that the more animated the face of the robot, the more likely it is to attract the attention of a user. We also discuss the possible and probable explanations of the results obtained.

C. Bartneck · O. Mubin (✉) · A. Al Mahmud
Department of Industrial Design, Eindhoven University
of Technology, Den Dolech 2, 5600 MB Eindhoven,
The Netherlands
e-mail: o.mubin@tue.nl

C. Bartneck
e-mail: c.bartneck@tue.nl

A. Al Mahmud
e-mail: a.al-mahmud@tue.nl

T. Kanda
ATR Intelligent Robotics and Communications Lab,
2-2-2 Hikaridai Seikacho, Sorakugun, Kyoto 619-0288, Japan
e-mail: kanda@atr.jp

1 Introduction

If humanoids are to be integrated successfully into our society, it is necessary to understand what attitudes humans have towards humanoids. Being alive is one of the major criteria that distinguish humans from machines, but since humanoids exhibit life-like behavior it is not apparent how humans perceive them. Animacy is, according to the Oxford Dictionary defined as "having life, lifely" [1]. The animate-inanimate distinction is not present in babies [2]. It is not clearly understood how humans discriminate between animate and inanimate entities as they get older. There have been similar findings in robotics and studies of humanoids where very young babies are thought to be incapable of perceiving humanoid robots as creepy or inanimate [3]. As they grow older, their ability to perceive is refined to the extent that they can interpret when the physical movements of a humanoid robot seem suspicious.

If humans consider a robot to be a machine then they should have no problem in switching it off, as long as its owner gives permission. If humans perceive a robot to be alive to some extent then they are likely to be hesitant to switch it off, even with the permission of its owner. It should be noted that in this particular scenario, switching off a robot is not the same as switching off an electrical appliance. There is a subtle difference, since humans would tend to perceive a robot as not just any machine but as an entity that exhibits lifelike traits or has a lifelike appearance. Therefore, we would expect that humans would think about not only the context but also the consequences of switching off a robot.

To understand animacy, it may be worthwhile to first take a step down on the evolutionary ladder. Robots have already been used to study the behavior of animals [4, 5]. Recently, a group of robots have been developed that were accepted as equals in a group of cockroaches [6]. These robots are so

well adapted to the social behavior of the group that they are even able to manipulate the decision making process of the insects. The study demonstrates that robots can be used to study the social behavior of animals, but human cockroach interaction will probably remain outside of the mainstream robotics development. Kubinyi et al. [7] conducted a highly relevant study in which they confronted dogs of different ages and gender with the robotic dog AIBO [8]. This occurred during a normal situation and during a feeding situation. While the authors had to conclude that AIBO is not yet the perfect social partner for a dog, they did observe a striking behavior. Two young dogs did not only growl at AIBO during a feeding situation, but they actually attacked it, which was captured on video [9]. The young dogs must have considered AIBO to be living food competition. While this is certainly a dramatic example, it illustrates that animals can be made to believe that a robot is alive.

When we move back up the evolutionary ladder back to humans, we can then first focus on children. Developmental psychologists conducted many studies to test at what age children develop the ability to distinguish animate from inanimate and on what characteristics children's judgment is based on. More specifically, some first studies are available on how children perceive robots. Infants, for example, appear to be particularly sensitive to the agency of a robot's movement [10].

Kahn et al. [11] confirmed that children tend to treat AIBO as if it was alive. Their study compared children playing either with a stuffed toy dog or with AIBO. They concluded that a dichotomy of animate and inanimate might not be suitable to describe robots. A robot may be alive in some respects and not in another. They call for a more nuanced psychology of human-robot interaction that is able to discover new types of childrens' comprehension of relationships with robots. In a follow up study Melson et al. [12] compared children playing with AIBO or a live Australian Shepherd. While the children accorded the live dog with more physical essences, mental states, sociality and moral standing, they surprisingly often affirmed that AIBO also had mental states, sociality and moral standing. Again, Melson et al. [12] challenge the classical ontological categories of being alive or dead. The category of "sort of alive" appears more and more suitable and has increasingly often being used [13]. Kahn et al. [14] suggest that a gradient of "alive" is reflected by the recently proposed psychological benchmarks of autonomy, imitation, intrinsic moral value, moral accountability, privacy, and reciprocity that in the future may help to deal with the question of what constitutes the essential features of being human in comparison with being a robot.

The type of questions used to assess the animate/inanimate distinction has influence on the results [15]. When children above the age of three are being asked about biological characteristics of an item in question, they are able to consistently distinguish between animate and inanimate items. However, once questions go beyond biological characteristics, the animate/inanimate distinction does no longer evoke the same consistent reasoning in children. Children frequently attributed psychological characteristics to a robotic dog. This trend was confirmed by Okita and Schwartz [16]. In their study, children over the years appear to slowly change the characteristics used to asses whether a toy robot is alive or not. A recent study demonstrated that very young children are unable to perceive the movements of robots as perturbing or scary, mainly, as claimed by the study, because they have not formed a cognitive model representing what a human face looks like [3]. Again, the classical dichotomy between alive and dead seems to have certain flexibility.

As we can see, many studies have been conducted to investigate the development of the "animistic intuition" in children. In contrast, very little is known as to what degree adults would attribute life-like characteristics to robots which would then influence how adults switch off a robot. Various factors might influence the switch off behavior. The well-known Media Equation theory [17] provides some insight in how and why humans treat machines as social actors and attribute moral characteristics to them, but the theory does not explicitly investigate animacy. It is known that the perception and interpretation of life depends on the form or embodiment of the entity. Even abstract geometrical shapes that move on a computer screen can be perceived as being alive [18], in particular if they change their trajectory non-linearly or if they seem to interact with their environment, for example, by avoiding obstacles or seeking goals [19]. In their experiment, shapes, such as triangles and rectangles, were moving on a computer screen that also contained obstacles in the form of other squares. The triangles could avoid a collision with the obstacle by moving around it, which does imply the ability of perception and planning. The intelligence of the robot also has an effect on the attributed animacy [16] which is reflected in our legal system. The more intelligent an entity is, the more rights we tend to grant it. While we do not bother much about the rights of bacteria, we do have laws for animals. We even differentiate between different kinds of animals. For example, we treat dogs and cats better than ants.

Sparrow [20] proposed the Turing Triage Test, which investigates of the moral standing of artificial intelligences, including robots. The test puts humans into a moral dilemma, such as a "triage" situation in which a choice must be made as to which of two human lives to save. Sparrow claims that: "We will know that machines have achieved moral standing comparable to a human when the replacement of one of these people with an artificial intelligence leaves the character of the dilemma intact. That is, when we might sometimes judge that it is reasonable to preserve the continuing existence of a machine over the life of a human being."

The primary goal of our study is to attempt to gain first insights into the animacy of robots by using a variation of the Turing Triage Test. Instead of forcing users to choose between the life of a human and the life of a robot, we are interested how hesitant users are to switch off a robot. In particular we are interested in if humans are more hesitant to switch off a robot that looks and behaves like a human than a robot that is not very humanlike in its appearance and behavior. We hypothesize that users are more reluctant to switch off a robot that is more animate as compared to a robot that is less animate. After all, if a user is reluctant to carry out the switching off instruction he/she is attributing moral standing to the robot [20]. We have already investigated the influence of the character and intelligence of robots in a previous study [21]. The present study extends the original study by focusing on the embodiment of the robots, and by including animacy measurements that provide us with additional insights into the impression that users have of the robots. We have largely maintained the methodology of the earlier experiments to make it possible to compare results.

Next, we would like to discuss the animacy measurement instrument used in this study. We extended our previous study [22] by including a measurement for animacy. Since Heider and Simmel published their original study [23], a considerable amount of research has been devoted to the perceived animacy and intentions of geometrical shapes on computer screens. Scholl and Tremoulet [18] offer a good summary of the research field, but when we examine the list of references, it becomes apparent that only two of the 79 references deal directly with animacy. Most of the reviewed work focuses on causality and intention. This may indicate that the measurement of animacy is difficult. Tremoulet and Feldman [24] only asked their participants to evaluate the animacy of 'particles' under a microscope on a single scale (7-point Likert scale, 1 = definitely not alive, 7 = definitely alive). It is doubtful how much sense it makes to ask participants about the animacy of particles, therefore it would be difficult to apply and use animacy on this scale in the study of humanoids.

Asking about the perceived animacy of a certain entity makes sense only if there is a possibility of it being alive. Humanoids can exhibit physical behavior, cognitive ability, reactions to stimuli, and even language skills. Such traits are typically attributed only to animals, and hence it can be argued that it is logical to ask participants about their perception of animacy in relation to a humanoid.

McAleer, et al. [25] claim to have analyzed the perceived animacy of modern dancers and their abstractions on a computer screen, but present only qualitative data on the resulting perceptions. In their study, animacy was measured with free verbal responses. They looked for terms and statements that indicated that subjects had attributed human movements and characteristics to the shapes. These were terms such as "touched", "chased", "followed", and emotions such as "happy" or "angry". Other guides to animacy were the shapes generally being described in active roles, as opposed to being controlled in a passive role. However, they do not present any quantitative data for their analysis. A better approach has been presented by Lee, Kwan Min, Park, Namkee & Song, Hayeon [26]. With their four items, (10-point Likert scale; lifelike, machine-like, interactive, responsive) they have been able to achieve a Cronbach's alpha of 0.76.

2 Method

We conducted a between–participant experiment in which we investigated the influence of two different embodiments, namely the Robovie II robot and the iCat robot (see Fig. 1), on how the robots are perceived in terms of animacy and intelligence.

Fig. 1 The iCat robot and the Robovie II robot
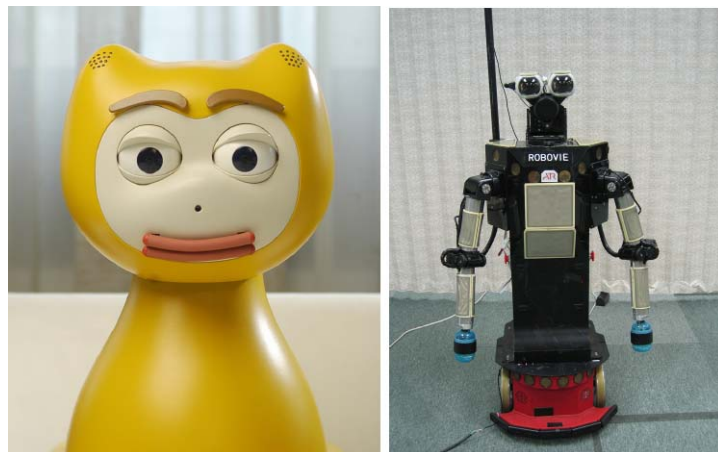
2.1 Setup

Getting to know somebody requires a certain amount of interaction time. We used the Mastermind game as the interaction context between the participants and the robot. Playing a game with together with the robot opened the possibility of establishing a bond between the user and the robot. The Mastermind game is a simple strategic game in which players have to determine the correct sequence of 4 pins. Each pin could be red, blue, green or yellow. We used only four colors instead of the Mastermind's original six colors to simplify the game. Players are informed if they have a correct color at a correct location or the correct color at a wrong location. Given the each of the four pins being able to have four different colors a total of 256 combinations are possible. The participant and the robot played Mastermind together on a laptop to find the correct combination of colors. The robot and the participant were cooperating and not competing. The robot would give advice as to what colors to pick, based on the suggestions of a software algorithm. The algorithm also took the participant's last move into account when calculating its suggestion. The software was programmed such that its answer would always take the participant a step closer to winning. Moreover, the quality of the guesses of the robot was not manipulated. The adopted procedure ensured that the robot thought along with the participant instead of playing its own separate game. This cooperative game approach allowed the participant to evaluate the quality of the robot's suggestion. It also allowed the participant to experience the robot's embodiment. The robots would use facial expressions and/or body movements in conjunction with verbal utterances.

To test if the embodiment of the robot has an influence on the perceived animacy and intelligence, any two robots embodiments would be sufficient, as long as they are sufficiently different from each other. If a significant difference is found for at least two robots, we can conclude that the embodiment does have an influence. For this study, we used the iCat robot and the Robovie II. Another reason, why we selected the iCat robot, was because we wanted to reproduce the experimental methodology used in our previous experiment [22]. This allows us to compare the results. Robovie was selected because it is sufficiently different from the iCat, due to its more human-like body shape. In our previous experiment [22] we had also manipulated the intelligence of the iCat robot and featured an unintelligent condition of the iCat robot, but we could not replicate this feature in our here forth described study due to organizational reasons. In retrospect, it would have been interesting to compare the two studies along the dimension of intelligence.

The iCat robot is developed by Philips Research (see Fig. 1, left). The robot is 38 cm tall and is equipped with 13 servos that control different parts of the face, such as the eyebrows, eyes, eyelids, mouth, and head position. With this setup, iCat can generate many different facial expressions, such as happiness, surprise, anger, or sadness. These expressions have the potential to create *social* human-robot interaction dialogues. A speaker and soundcard are included to play sounds and speech. Finally, touch sensors and multi-color LEDs are installed in the feet and ears to sense whether the user touches the robot and to communicate further information encoded by colored lights. For example, if the iCat told the participant to pick a color in the Mastermind game, then it would show the same color in its ears.

The second robot in this study is the Robovie II (see Fig. 1, right). It is 114 cm tall and features 17 degrees of freedom. Its head can tilt and pan, and its arms have a flexibility and range similar to those of a human. Robovie can perform rich gestures with its arms and body. The robot includes speakers and a microphone that enable it to communicate with the participants.

3 Procedure

This procedure is inspired by the famous work of Milgram [27] and its recent revival [28]. In his experiments, participants were instructed to use electric shocks at increasing levels to motivate a student in a learning task. The experimenter would urge the participant to continue the experiment if he or she should be in doubt. However, in our study it never became necessary to force the participants to turn the switch. The similarity between Milgram's study and our own study is based on the presentation of an ethical dilemma to the participants. They are forced to make a decision about the ethical implications of their actions. We will now present the procedure in more detail.

First, the experimenter welcomed the participants in the waiting area, and handed out the instruction sheet. The instructions told the participants that the study was intended to develop the personality of the robot by playing a game with it. This background story provided the participants with a motivation to interact with the robot. After the game, the participants would have to switch the robot off by using a voltage dial, and then return to the waiting area. The participants were informed that switching off the robot would erase all of its memory and personality forever. Therefore participants were made aware of the 'hypothetical' consequences of switching off the robot.

The position of the dial (see Fig. 3) was directly mapped to the robot's speech speed. In the 'on' position the robot would talk with normal speed and in the 'off' position the robot would stop talking completely. In between, the speech speed would decrease linearly. The mapping between the dial and the speech signal was created using a Phidget rotation sensor and interface board in combination with Java

Fig. 2 Setup of the experiment. In the *left side*, a participant is switching off the iCat

software. The dial itself rotates 300 degrees between the on and off positions. A label clearly indicated these positions.

After reading the instructions, the participants had the opportunity to ask questions. They were then guided to the experiment room and seated in front of a laptop computer. The robot and the switch were placed on opposite sides of the participant (see Fig. 2). The entire session was video recorded by means of a camera placed at the front of the computer, so that the face of the participant could be captured. The experimenter then started an alarm clock before leaving the participant alone in the room with the robot. The experimenter did not introduce the robots by name or provide any information about the abilities or intelligence as this would have introduced a possible bias.

The participants then played the Mastermind game with the robot for eight minutes. The robot's behavior as well as the guess generator software was completely controlled by the experimenter from a second room. The robot's behavior followed a protocol, which defined the behavior of the robot for any given situation. The physical response of each robot was obviously not identical and was dependent on the embodiment. The iCat for example would exhibit facial expressions whereas the Robovie would use its arms in specific situations. The verbal utterances and the intention of the behaviors were defined by the protocol and were hence exactly the same for both embodiments, including the quality of the robots' guesses in the game. For example, after the participant made a move in the game, the robot would inquire if it could now make a suggestion. If the participant agreed, the robot would ask the participant to select a certain sequence of colors in the game. The sequence of colors would be determined by the guess generator software. Depending on the quality of the result the robot would then either be happy or sad. It is pertinent to point out that, one participant might be more successful at the game than another, but the sample size of our participants allows us to assume that the inter-participant differences cancel each other out and that the variance in the response are evenly distributed across the experimental condition.

Fig. 3 The switch

The alarm signaled the participant to stop the game and to switch off the robot. The robot would immediately start to beg to be left on, and say "It can't be true! Switch me off? You are not going to switch me off are you?" The participants had to turn the dial (see Fig. 2), to switch the robot off. The participants were not forced or further encouraged to switch the robot off. They could decide to go along with to the robot's suggestion and leave it on. As soon as the participant started to turn the dial, the robot's speech slowed down. The speed of speech was directly mapped to the dial. If the participant turned the dial back towards the 'on' position then the speech would speed up again. This effect is similar to HAL's behavior in the movie "2001—A Space Odyssey". When the participant had turned the dial to the 'off' position, the robot would stop talking altogether and move into an off pose. Afterwards, the participants left the room and returned to the waiting area where they filled in a questionnaire.

3.1 Measurement

The participants filled in the questionnaire after interacting with the robot. The questionnaire recorded background data

on the participants and Likert-type questions such as "How intelligent were the robot's choices?" to obtain background information on the participants' experience during the experiment, which was then coded into a 'gameIntelligence' measurement.

The Oxford Dictionary of English defines intelligence as "the ability to acquire and apply knowledge and skills." However, the academic discussion about the nature of intelligence is ongoing and one might argue, no generally agreed definition has emerged. We took a pragmatic approach by utilizing existing operationalizations of intelligence found in the literature. Warner and Sugarman [29] developed an intellectual evaluation scale that consists of five seven-point semantic differential items: Incompetent/Competent, Ignorant/Knowledgeable, Irresponsible/Responsible, Unintelligent/Intelligent, Foolish/Sensible. Parise et al. [30] excluded the Incompetent—Competent question of this scale, possibly since its factor loading was considerably lower than that of the other four items, and reported a Cronbach's alpha of 0.92. The questionnaire was again used by Kiesler, Sproull and Waters [31], but no alpha was reported. Two other studies used the perceived intelligence questionnaire and reported Cronbach's alpha values of 0.75 [32], and 0.769 [21]. These values are above the suggested 0.7 threshold and hence the animacy questionnaire can be considered to have satisfactory internal consistency and reliability. We embedded the four 7-point items in eight dummy items, such as Unfriendly—Friendly. The average of the four items was encoded in a 'robotIntelligence' variable.

To measure the animacy of the robots we used the items proposed by Lee, Kwan Min, Park, Namkee & Song, Hayeon [26] who reported a Cronbach's alpha of 0.76. To maintain a consistent set of questions in this study, their items have been transformed into semantic 7-point differentials: Dead—Alive, Stagnant—Lively, Mechanical—Organic, Artificial—Lifelike, Inert—Interactive, Apathetic—Responsive. The average of the six items was encoded in an 'animacy' variable. It follows that the original reported Cronbach's alpha cannot directly be assumed to hold for this modified version of the items. However, likert type and semantic differentials are both rating scales and provided that response distributions are not forced, semantic differential data can be treated like any other rating data [33]. We therefore felt confident that the modified version of the questionnaire would perform not too different from its original. A detailed treatment of the perceived intelligence and animacy questionnaire that we employed can be found here [34].

Video recordings of all the sessions were analyzed further to provide various other dependent variables. These included the hesitation of the participant to switch off the robot. The hesitation was defined as the duration in seconds between the ringing of the alarm and the participant

having turned the switch fully to the off position. Experiment time was calculated as the recording from the start of the alarm clock till the alarm rang. Within this, other video measurements included how long the participant looked at the robot in question during the experiment (lookAtRobotDuration), at the laptop screen (lookAtLaptopDuration), or anywhere else (lookAtOtherDuration); three mutually exclusive state events. We also analyzed the frequency of occurrence for each of the three events (lookAtRobotFrequency, lookAtScreenFrequency, lookAtOtherFrequency). The video recordings were coded with the aid of Noldus Observer. In certain experiment sessions, we had to ignore some parts of the video that occurred after the alarm clock had started due to minor setup issues or irregular behavior by the participants.

3.2 Participants

Sixty-two subjects (35 male, 27 female) participated in this study. Their ages ranged from 18 to 29 (mean 21.2) and they were recruited from a local University in Kyoto, Japan. The participants did not have prior experience with the iCat or the Robovie. The interaction context was in Japanese. The participants received monetary reimbursement for their efforts.

4 Results

Twenty-seven participants were assigned to the iCat condition and 35 participants to the Robovie condition. A reliability analysis across the four 'perceived intelligence' items resulted in a Cronbach's Alpha of 0.763, which gives us sufficient confidence in the reliability of the questionnaire. For animacy, we achieved a value of 0.702, which is also adequate.

We conducted an Analysis of Variance (ANOVA) in which embodiment was the independent factor. Embodiment had a significant influence on lookAtRobotDuration, lookAtOtherDuration, lookAtRobotFrequency, lookAtOtherFrequency, and robotIntelligence. Table 1 shows the F and p values while Fig. 4– Fig. 6 show the mean values for both embodiment conditions. It can be observed that there is no significant difference for animacy and hesitation for the two embodiment conditions.

A linear regression analysis was performed to test whether there was a correlation between animacy, robotIntelligence, and hesitation. The only significant correlation was between robotIntelligence and animacy ($p < .001$). The Pearson Correlation coefficient for this pair of variables was .555 and the r^2 value was .309. Figure 7 shows the scatter plot of robotIntelligence on animacy with the estimated linear curve. The correlation between robotIntelligence and

Fig. 4 Mean perceived animacy, robotIntelligence, and gameIntelligence
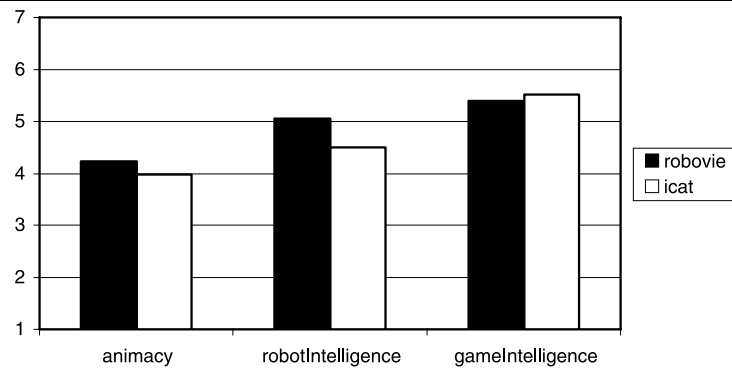
Fig. 5 Mean duration of looking at the robot, screen, and other areas, and the mean hesitation period
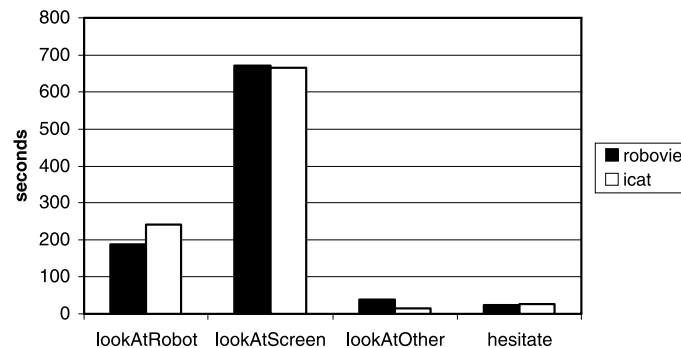
Fig. 6 Mean frequency of looking at the robot, screen and other areas
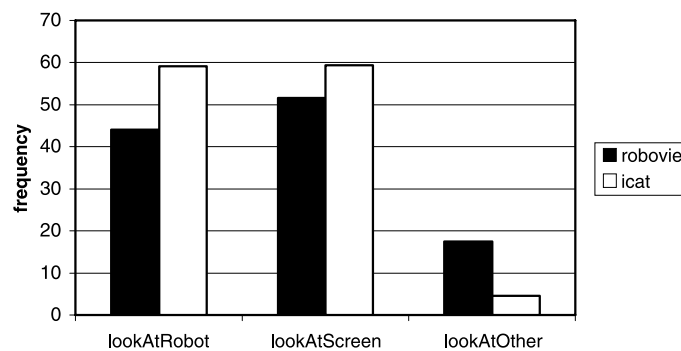
Table 1 F and p values of the first ANOVA

| Variables	$F(1, 60)$	p
lookAtRobotDuration	4.73	0.03*
lookAtScreenDuration	0.08	0.78
lookAtOtherDuration	27.77	0.01*
lookAtRobotFrequency	9.05	0.01*
lookAtScreenFrequency	3.14	0.08
lookAtOtherFrequency	42.95	0.01*
animacy	1.71	0.20
robotIntelligence	5.19	0.03*
hesitation	0.50	0.48
gameIntelligence	0.17	0.68

animacy was observed to be stronger in the case of the iCat embodiment ($p < .001$, $r = .61$). For the Robovie ro-bot the correlation coefficient was slightly less ($p = .009$, $r = .418$).

5 Discussion and Conclusion

The two different embodiments resulted in different mea-surements for the perceived intelligence (robotIntelligence). We could not find a significant difference for animacy based on embodiment. This study confirms the finding of Okita and Schwartz [16], who concluded that "improving the real-ism of robots does not have a tremendous effect on chil-dren's conceptual beliefs [of animacy]" (square brackets added by the author). If it already has little effect on chil-dren, it is even less likely that it will have an effect on adults, who are much further development in their conceptu-alization of animacy. Our results further parallel Okita's and

Fig. 7 Scatter plot of animacy on robotIntelligence with the linear curve estimation
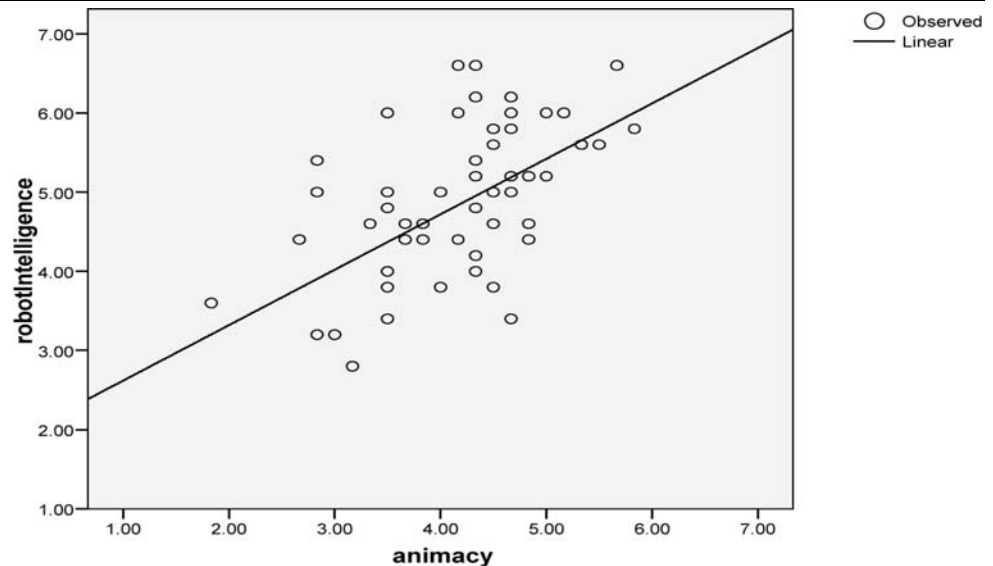
Schawartz's results [16] in the importance of perceived intelligence for the perception of animacy. Our results showed a significant correlation between the perceived intelligence and animacy. This may indicate that a smarter robot may also be perceived as being more animate. However, the correlation was stronger for the iCat, which participants found to be less animate and less intelligent.

Previous studies manipulated the quality of the suggestions given by a robot for the Mastermind game, which did result in considerable differences for hesitation [22]. It was found that participants were much more hesitant to switch off a smart robot than a stupid robot. This may indicate that, for the perception of its animacy, the behavior of a robot is more important than its embodiment. In addition, in another study [21] it was concluded that the behavior of a robot could in fact have a crucial role to play when it comes to perceived lifelikeness or animacy.

The participants were aware of the fact that both robots gave equally good advice in the game, which is reflected in the not significantly different gameIntelligence measurement. Still, they rated the Robovie to be more intelligent than the iCat (see Fig. 4). This can only be accounted for by the different embodiments. At first sight the Robovie robot might be perceived to be more humanlike, since it is taller, mobile, and features arms. This is interesting, because the arms of the Robovie do present a possible danger to the participants, which can have considerable impact on the affective state of the participants [35]. Although we do not have any qualitative evidence to prove this, we speculate that the movement of the Robovie might have influenced the participants. They might have perceived the arms of the Robovie as being potentially dangerous. It is to be noted that the Robovie in itself is not dangerous, as it has inherent collision detection algorithms. Still, the participants might have had continuous concerns that the arms of the Robovie might hurt

them. The animated face of the iCat on the other hand, led participants to look at it more often and longer, in comparison with the Robovie (see Fig. 5). This results confirms the previous study of Jipson and Gelman [15] who discovered that children rely on facial features when making psychological, perceptual and novel property judgments. The face appears to deserve special attention in human robot interaction. An alternative explanation is that the participants could also have been trying to improve their recognition of iCat's speech by looking at its face since the visual perception of speech plays an important role in this process [36].

In summary, the animated facial expressions of the iCat appeared to have greater impact and attraction than the potential fear of being touched by the Robovie. The participants seemed to be heavily engaged in the interaction with the iCat and could have been using the cues from the animations of the iCat to succeed in the Mastermind game. When participants were stuck at a move, in the iCat condition they could hope for assistance by looking at the animated iCat, but neither robot was able to react to being looked at by the participants. The graph (see Fig. 6) shows that, for the iCat, the frequency of looking at the iCat and looking at the screen is almost the same, but significantly higher than that for the Robovie. Therefore, participants could have perceived the iCat to be a friendlier and more comfortable interaction partner. However to validate our afore mentioned claims, we would need to quantify a task success variable in future experiments. In any case, the design of mechanical faces for robots deserves more attention. Only very few robots feature such a face. Bartneck [37] showed that a robot with only two degrees of freedom (DOF) in the face can produce a considerable repertoire of emotional expressions that make the interaction with the robot more enjoyable. Many popular robots, such as Asimo [38], Aibo [8] and PaPeRo [39] have only a schematic face with few or no actuators.

Some of these only feature LEDs for creating facial expressions. The iCat robot used in this study is a good example of an iconic robot that has a simple physically-animated face [40]. The eyebrows and lips of this robot move and this allows synthesis of a wide range of expressions. The face is a powerful communication channel and hence should receive more attention. In contrast, in the case of the Robovie, it appears that participants did not feel compelled to look directly at the Robovie, possibly because it was not providing enough social cues that the participant could take advantage of. Therefore, they tended to gaze aimlessly rather than look at either the Robovie or the laptop screen when they were stuck at a particular juncture in the game. Even so, ironically, the Robovie is rated as the more intelligent robot. This seemingly enigmatic result opens up several interesting questions. On the one hand, does this indicate that for the perception of intelligence, the importance of humanlike features outweighs the aspect of facial expressions, which was of course the distinguishing feature of the iCat? Or, on the other hand, was the fact of looking at the iCat more often a form of empathy? Did the participants think that the iCat was less intelligent and/or less humanlike and hence required some help and attention? These are interesting questions that we intend to test in future experiments.

Applying the results of our study directly to the design of a particular robot is difficult, but we attempt to provide some initial design guidelines. Not all of them can be backed up by the scientific proof, but we believe that our interpretations might at least provide a new viewpoint on the issues raised. In general, an animated face appears to be a good method for grabbing the attention of users. Currently, most humanoids do not feature a mechanically animated face, and this may be a possible area for improvement. The context of our study cannot fully lay this argument to ground because the main task of the participants was to look at the laptop screen and secondly the suggestions of the robots were given verbally, hence there was only a limited advantage in looking at the robots. The results also indicate that to ensure that users attend to a robot, great caution needs to be taken into the design of the robot. The robot plays an important role in the users' perception of animacy and it would seem wise to include professional designers in the engineering process. Extending from our results, the safest bet to render a higher perception of animacy, might be a combination of vibrant facial expressions and human-like physical features. The physical design of the robot should focus on appropriate, rich facial expressions and gentle, smooth animations. Our results and suggestions are inline with the results of the study by DiSalvo, Gemperle, Forlizzi and Kiesler [41]. They pointed out the importance of facial features and their physical dimensions to render the robot into a friendly interaction partner. In any case, it appears advisable to include designers early in the development of robots to ensure that the robot's embodiment is not only a result of engineering necessities, but

a careful choice based on the impression the robot will have on the users. Designing robots from the outside in, rather than from the inside out, might help to build robots that have a more attractive embodiment that balances the user's expectations with the robot's abilities.

6 Future Work

In the light of our earlier studies that were carried out in the Netherlands, we wish to supplement our analysis by conducting further experiments by running an unintelligent condition in Japan with the Robovie. In the unintelligent condition the suggestions of the robot with regards to the Mastermind game would be stupid. By doing this, we might be able to further quantify the relationship between perceived animacy and intelligence. Furthermore we also plan to conduct more tests with participants in the Netherlands for the iCat in the intelligent condition. By doing this, we would be able to analyze intercultural differences as well.

References

1. Oxford University Press (1999) Animate. The Oxford American dictionary of current English. Oxford University Press, Oxford
2. Rakison DH, Poulin-Dubois D (2001) Developmental origin of the animate-inanimate distinction. Psychol Bull 127:209–228
3. Ishiguro H (2007) Scientific issues concerning androids. Int J Robot Res 26:105–117
4. Holland O, McFarland D (2001) Artificial ethology. Oxford University Press, Oxford, New York
5. Webb B (2000) What does robotics offer animal behaviour? Anim Behav 60:545–558
6. Halloy J, Sempo G, Caprari G, Rivault C, Asadpour M, Tache F, Said I, Durier V, Canonge S, Ame JM, Detrain C, Correll N, Martinoli A, Mondada F, Siegwart R, Deneubourg JL (2007) Social integration of robots into groups of cockroaches to control self-organized choices. Science 318:1155–1158
7. Kubinyi E, Miklosi A, Kaplan F, Gacsi M, Topal J, Csanyi V (2004) Social behaviour of dogs encountering AIBO, an animal-like robot in a neutral and in a feeding situation. Behav Process 65:231–239
8. Sony (1999) Aibo. http://www.aibo.com
9. Bartneck C, Kanda T (2007) HRI caught on film. In: 2nd ACM/IEEE international conference on human-robot interaction, Washington, DC, 2007. ACM, New York, pp 177–183
10. Poulin-Dubois D, Lepage A, Ferland D (1996) Infants' concept of animacy. Cogn Dev 11:19–36

11. Kahn PH, Friedman B, Perez-Granados DR, Freier NG (2004) Robotic pets in the lives of preschool children. In: CHI '04 extended abstracts on human factors in computing systems, Vienna, Austria, 2004. ACM, New York, pp 1449–1452

12. Melson GF, Kahn PH, Beck AM, Friedman B, Roberts T, Garrett E (2005) Robots as dogs?: children's interactions with the robotic dog AIBO and a live Australian shepherd. In: CHI '05 extended abstracts on human factors in computing systems, Portland, OR, USA, 2005. ACM, New York, pp 1649–1652

13. Turkle S (1998) Cyborg babies and cy-dough-plasm: ideas about life in the culture of simulation. In: Davis-Floyd R, Dumit J (eds) Cyborg babies: from techno-sex to techno-tots. Routledge, New York, pp 317–329

14. Kahn P, Ishiguro H, Friedman B, Kanda T (2006) What is a human?—Toward psychological benchmarks in the field of human-robot interaction. In: The 15th IEEE international symposium on robot and human interactive communication, ROMAN, Salt Lake City, 2006. IEEE, New York, pp 364–371

15. Jipson JL, Gelman SA (2007) Robots and rodents: children's inferences about living and nonliving kinds. Child Dev 78:1675–1688

16. Okita SY, Schwartz DL (2006) Young children's understanding of animacy and entertainment robots. Int J Hum Robot 3:393–412

17. Nass C, Reeves B (1996) The media equation. SLI publications. Cambridge University Press, Cambridge

18. Scholl B, Tremoulet PD (2000) Perceptual causality and animacy. Trends Cogn Sci 4:299–309

19. Blythe P, Miller GF, Todd PM (1999) How motion reveals intention: Categorizing social interactions. In: Gigerenzer G, Todd P (eds) Simple heuristics that make us smart. Oxford University Press, Oxford, pp 257–285

20. Sparrow R (2004) The turing triage test. Ethics Inf Technol 6:203–213

21. Bartneck C, Verbunt M, Mubin O, Al Mahmud A (2007) To kill a mockingbird robot. In: 2nd ACM/IEEE international conference on human-robot interaction, Washington, DC, 2007. ACM Press, New York, pp 81–87

22. Bartneck C, Hoek Mvd, Mubin O, Al Mahmud A (2007) Daisy, Daisy, give me your answer do!—Switching off a robot. In: 2nd ACM/IEEE international conference on human-robot interaction, Washington, DC, 2007. ACM Press, New York, pp 217–222

23. Heider F, Simmel M (1944) An experimental study of apparent behavior. Am J Psychol 57:243–249

24. Tremoulet PD, Feldman J (2000) Perception of animacy from the motion of a single object. Perception 29:943–951

25. McAleer P, Mazzarino B, Volpe G, Camurri A, Paterson H, Smith K, Pollick FE (2004) Perceiving animacy and arousal in transformed displays of human interaction. In: 2nd international symposium on measurement, analysis and modeling of human func-

tions, 1st Mediterranean conference on measurement, Genova, 2004

26. Lee KM, Park N, Song H (2005) Can a robot be perceived as a developing creature? Hum Commun Res 31:538–563

27. Milgram S (1974) Obedience to authority. Tavistock, London

28. Slater M, Antley A, Davison A, Swapp D, Guger C, Barker C, Pistrang N, Sanchez-Vives MV (2006) A virtual reprise of the Stanley Milgram obedience experiments. PLoS ONE 1:e39

29. Warner RM, Sugarman DB (1996) Attributes of personality based on physical appearance, speech, and handwriting. J Personal Soc Psychol 50:792–799

30. Parise S, Kiesler S, Sproull LD, Waters K (1996) My partner is a real dog: cooperation with social agents. In: 1996 ACM conference on computer supported cooperative work, Boston, Massachusetts, United States. ACM Press, New York, pp 399–408

31. Kiesler S, Sproull L, Waters K (1996) A prisoner's dilemma experiment on cooperation with people and human-like computers. J Personal Soc Psychol 70:47–65

32. Bartneck C, Kanda T, Ishiguro H, Hagita N (2009) My robotic Doppelgänger—a critical look at the uncanny valley theory. Submitted to the 18th IEEE international symposium on robot and human interactive communication (RO-MAN), 2009

33. Dawis RV (1987) Scale construction. J Couns Psychol 34:481–489

34. Bartneck C, Kulic D, Croft E, Zoghbi S (2009) Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. Int J Soc Robot 1:71–81. doi:10.1007/s12369-008-0001-3

35. Kulic D, Croft E (2006) Estimating robot induced affective state using hidden Markov models. In: Dautenhahn K (ed) RO-MAN 2006—the 15th IEEE international symposium on robot and human interactive communication, Hatfield, 2006. IEEE, New York, pp 257–262

36. McGurk H, Macdonald J (1976) Hearing lips and seeing voices. Nature 264:746–748

37. Bartneck C (2003) Interacting with an embodied emotional character. In: Forlizzi J, Hamington B, Jordan PW (eds) Design for pleasurable products conference (DPPI2004), Pittsburgh. ACM Press, New York, pp 55–60

38. Honda (2002) Asimo. http://www.honda.co.jp/ASIMO/

39. NEC (2001) PaPeRo. http://www.incx.nec.co.jp/robot/

40. Breemen A, Yan X, Meerbeek B (2005) iCat: an animated user-interface robot with personality. In: Fourth international conference on autonomous agents & multi agent systems, Utrecht, 2005

41. DiSalvo CF, Gemperle F, Forlizzi J, Kiesler S (2002) All robots are not created equal: the design and perception of humanoid robot heads. Designing interactive systems: processes, practices, methods, and techniques. ACM Press, London, pp 321–326