Evaluating the Effect of Saliency Detection and Attention Manipulation in Human-Robot Interaction

Guido Schillaci · Saša Bodiroža · Verena Vanessa Hafner

Abstract The ability to share the attention with another individual is essential for having intuitive interaction. Two relatively simple, but important prerequisites for this, saliency detection and attention manipulation by the robot, are identified in the first part of the paper. By creating a saliency based attentional model combined with a robot ego-sphere and by adopting attention manipulation skills, the robot can engage in an interaction with a human and start an interaction game including objects as a first step towards a joint attention.

We set up an interaction experiment in which participants could physically interact with a humanoid robot equipped with mechanisms for saliency detection and attention manipulation. We tested our implementation in four combinations of activated parts of the attention system, which resulted in four different behaviours.

Our aim was to identify those physical and behavioural characteristics that need to be emphasised when implementing attentive mechanisms in robots, and to measure the user experience when interacting with a robot equipped with attentive mechanisms.

We adopted two techniques for evaluating saliency detection and attention manipulation mechanisms in human-robot interaction: user experience as measured by qualitative and quantitative questions in questionnaires and proxemics estimated from recorded videos of the interactions.

The robot's level of interactiveness has been found to be positively correlated with user experience factors like excitement and robot factors like lifelikeness and intelligence, suggesting that robots must give as much feedback as possible in order to increase the intuitiveness of the interaction, even when performing only attentive behaviours. This was confirmed also by proxemics analysis: participants reacted more frenetically when the interaction was perceived as less satisfying. Improving the robot's feedback capability could increase user satisfaction and decrease the probability of unexpected or incomprehensible user movements. Finally, multi-modal interaction (through arm and head movements) increased the level of interactiveness perceived by participants. Positive correlation has been found between the elegance of robot movements and user satisfaction.

Keywords Measuring interaction · Attentional models · Multimodal interaction · Human-robot interaction · Proxemics

G. Schillaci (✉) · S. Bodiroža · V.V. Hafner
Cognitive Robotics Group, Department of Computer Science, Humboldt-Universität zu Berlin, Berlin, Germany
e-mail: guido.schillaci@informatik.hu-berlin.de

S. Bodiroža
e-mail: bodiroza@informatik.hu-berlin.de

V.V. Hafner
e-mail: hafner@informatik.hu-berlin.de

1 Introduction

Current social robotic systems require interaction protocols which decrease the intuitiveness of the interaction itself, causing frustration and despair in the user. Recently, interest has been focused on measuring the efficacy of robot behaviours and its perceived intelligence based on the evaluation from human users [1]. Indeed, measuring human-robot interaction could suggest what to improve in the cognitive abilities and in the appearance of the robot and how to improve it.

When human-robot interaction fails, the reason most often lies in the fact that the robot and the human try to communicate about different things and that the human partner has wrong expectations of the robotic partner. Several prerequisites have been identified [2, 3] about the features (both physical and cognitive) that let a robot interact effectively and naturally with a human user.

Here, we stress the fact that robots need to reach joint attention with the users for having successful interactions. This has not been achieved so far, since joint attention not only requires visual attention on the same visual features in the environment, but also skills in attention detection, attention manipulation, social interaction skills and even intentional understanding [2]. Without joint attention a robot will not be able to achieve a degree of interaction comparable to a human-human interaction.

Previously, we implemented an attentive mechanism which adopts two fundamental skills for joint attention [4]. In this paper, we focus on measuring the quality of this implementation. By evaluating robot skills, in fact, we want to identify those characteristics that need to be emphasised when implementing attentive mechanisms in robots and to identify correlations between them.

Several metrics for measuring HRI have been proposed, from measuring the ability of a robot to engage in temporally structured behavioural interactions with humans [5], to evaluating robot social effectiveness from different points of view (engineering, psychological, sociological) [6]. We adopted a series of metrics based on cognitive science studies about measuring social skills in humans and based on studies about how robots are perceived by humans and whether this perception affects the expectation humans have about robot intelligence (Godspeed questionnaire [7]).

Quantifying human behaviour usually requires the analysis of video recordings, questionnaires and interviews. In this work, we used the first two methods for quantifying the quality of robot behaviour. We set up four interaction experiments between a humanoid robot and a user and recorded them. After each experiment, the user was asked to fill a questionnaire on the quality of the interaction and on the perception of several functional and physical properties of the robot. To the best of our knowledge, very few studies have been done so far on correlating human perception of robot skills (measured with the Godspeed questionnaire, whose reliability was tested) with proxemic distances.

In [8], Takayama and Pantofaru adopted part of the Godspeed questionnaire in their measurements, finding that people who held more negative attitudes toward robots felt less safe when interacting with them. They also studied human personal space around robots, finding that experience with owning or experience with robots decreases the personal space that people maintain around robots and a robot looking people in the face influences proxemic behaviours. The latter suggests to perform proxemics analysis when measuring attentive mechanisms in robots.

The article is organised as follows: Section 2 introduces the saliency detection and attention manipulation skills implemented on the Nao robot from Aldebaran; Section 3 shows the experimental setup: experimental procedure, the robot platform, structure of the participants, measurements performed, results and discussion; finally, in Sect. 4 we depict the achievements of the current work and how we would like to continue it.

2 Saliency Detection and Attention Manipulation

In this section, we will provide a short overview of the system we implemented on the humanoid robot Nao by which we provide the robot with both saliency detection and attention manipulation skills [9]. For full description and an overview of work in this area, please see [4].

Attention is a cognitive skill, studied in humans and observed in some animal species, which lets a subject concentrate on a particular aspect of the environment without the interference of the surrounding. There is evidence from developmental psychology studies that the development of skills to understand, manipulate and coordinate attentional behaviour lays the foundation of imitation learning and social cognition [10].

In our world, we are constantly surrounded with items, such as objects, people and events, which stand out to their neighbouring items. This is represented with the saliency of those items. Saliency detection represents an attentional mechanism, through which those items are discovered, and it enables humans to shift their limited attentional resources to those objects that stand out the most.

There are two approaches that can be combined—a bottom-up, pre-attentive process and a top-down process influenced by motivation. Bottom-up detection uses different low-level features (e.g. motion, colour, orientation and intensity) for saliency detection. Top-down detection relies on high-level features, and it is highly influenced by our current goals and intentions. The combination of bottom-up and top-down processes is highly inspired by similar mechanisms in humans [11, 12].

Figure 1 gives an overview of the attention mechanism we implemented on the humanoid robot Nao.

For saliency detection, we used optic flow and face detection filters that store the information in a robot ego-sphere, and a marker detector for simplified object detection. Each feature detector represents one filter, and by applying it to the input, a saliency map is generated. The robot directs its attention to the point which has the highest saliency. Due to Nao's computational limitations, the ego-sphere is represented with a tessellated sphere, where information about
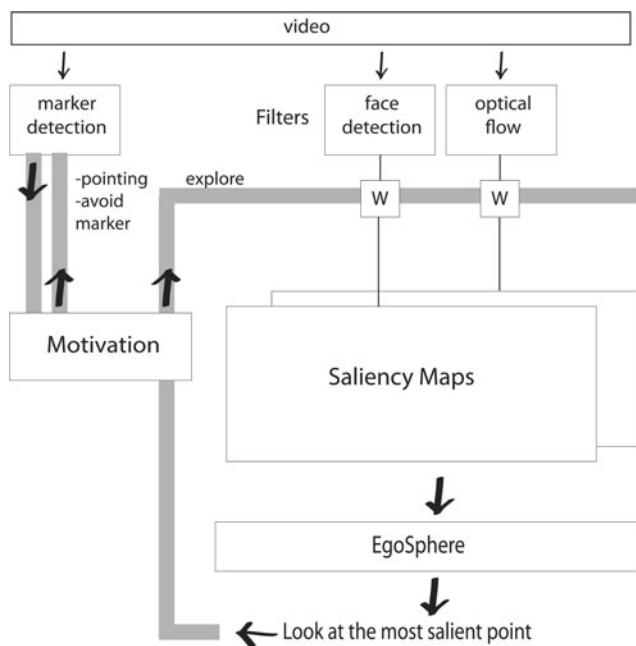
Fig. 1 Overview of the attentive mechanism. Frames are analysed by three different filters which are activated by the motivation system. Optic flow and face detection filters feed the ego-sphere, while the marker detector filter stores objects in a different memory. The motivation system activates or deactivates filters and movements according to its current state. See Sect. 3 and refer to [4] for further information

salient areas is stored in the edges of the sphere, like in [13, 14]. To simulate a short-term memory, habituation, inhibition and decay mechanisms are employed [15].

Pointing is a way for manipulating the attention of someone else. It is still not clear whether this behaviour is innate or if it results from reaching behaviours in its first developmental stage. Recognising and performing pointing gestures is very important for being able to share attention with another person [2].

We implemented learning through self-exploration on a humanoid platform [16]. We used motor babbling for learning the mapping between different sensory modalities and for equipping the robot with prediction abilities of sensory consequences (in this case, the position of the hand of the robot) from control commands applied to its neck and its arm [3]. Then, we equipped the robot with prediction abilities of arm movement commands that allowed for and resulted in pointing towards an object presented outside the reach of the robot [9].

Finally, we implemented a partially preprogrammed motivation system to show how different behaviours can result in the activation or deactivation of parts of the attention system, actually implementing a top-down approach for saliency detection, or in the activation of attention manipulation.

3 Experiment

The proposed experiment aimed at several goals: test the quality of the implemented saliency detection and attention manipulation mechanisms; identify those physical and behavioural characteristics that need to be emphasised when implementing attentive mechanisms in robots; measure the user experience when interacting with a robot equipped with attentive mechanisms; find correlations between heterogeneous robot features perceived by the participants during the exhibition of attentive mechanisms; and analyse the differences in the perception depending on the different behaviours performed by the robot.

We tested our implementations in four combinations of activated parts of the attention system, which resulted in four different behaviours:

Exploration. In this state, the robot is attracted by movements, faces and objects, actually looking like exploring the surrounding environment.

Interaction. This behaviour reproduces the experiment done in [9]. The robot is looking and pointing at an object, if there is one.

Interaction avoidance. This behaviour implements the loss of interest and boredom. In this state the robot looks away from the object handed over by the interacting partner.

Full interaction. This behaviour is composed as a sequence of the previous behaviours. The first performed action is *exploration*. Once the robot has detected a person to interact with and an object which can be used to draw the attention of the user, its motivation state changes to *interaction*, and after a certain period it switches to *interaction avoidance*, which is followed by *exploration*.

For a full description of the behaviours, please refer to [4].

3.1 Hypotheses

We had several expectations about the outcomes of the experiment. We expected that the level of interactiveness of the robot was positively correlated with the level of excitement and perceived intelligence. Playing with the robot in the interaction state might be more exciting and satisfactory than playing with it in the avoid interaction one.

Multi-modal interaction (through arms and head movements) might increase the perception of interactiveness; on the other side, a less interactive behaviour might decrease user satisfaction and cause the participants to behave nervously.

Anthropomorphic attributes might be positively correlated with the perception of intelligence.

Reaching commands can be perceived as a desire to grasp the object. This has been demonstrated in a preliminary ex-

Fig. 2 Experimental setup showing interaction between the Nao and a person

periment, in which the participants were asked how they interpreted the movements of the robot performing the interaction behaviour.

3.2 Procedure

The experiments consisted of the robot performing the behaviours described in the previous section in four separate interaction sessions, one per each of the four behaviours. The experiment supervisor manually activated or deactivated them. Figure 2 shows a frame taken from a typical interaction session. The user sat in front of the robot at a distance of ca. 90 cm. For each person, each interaction test lasted one minute. We recorded the interaction with a standard camera (resolution 640 × 480) placed at ca. 2 meters perpendicularly to the robot-user axis. Beside the table where the robot was standing there was a scale drawn on a whiteboard for the visual estimation (estimated average error: 5 cm) of the distance between the nose of the user and the head of the robot and from the hand of the user and the head of the robot; according to the type of interaction, we noticed that the users move their hands closer to the robot.

After each of the four interaction sessions, the participants were asked to fill a questionnaire about the quality of the interaction with the robot and about the perception of robot behaviours.

3.3 Robot Platform

The robot platform is the Nao (version 3.3) from Aldebaran, a humanoid robot around 57 cm tall. For the experiment, we used only the degrees of freedom in the arms and the neck. The lower camera is positioned below two eyes, which resulted in the robot not seeing an object if it is brought close to the eyes. For that reason two fake eyes were placed on the sides of the lower camera, and the real eyes were covered with a tape.

The attention mechanism was implemented in C++ using the framework of the Nao Team Humboldt [17]. The attentional mechanism is fully executed onboard the robot and there is no remote processing of the data. The robot is connected to the computer through Ethernet. A robot control program is running on the computer which is used to visualise the data and activate required modules for the attentional mechanism in the framework.

We adopted such a robot for measuring the users' expectations about the robot's skills due to its anthropomorphic form. Moreover, its small child-resembling size could reduce users' expectations, thus increasing the positive evaluation of the interactions.

Unfortunately, Nao has limited computational resources. Our implementation, at the current state, lets the robot process all the filters at a rate of approximately 7–8 frames per second. The computationally most expensive algorithms are those related to image processing, e.g. the face detection filter and the optical flow filter, which together take almost 110 ms per calculation. This results in slower movements and reactions when the robot is in the exploration state and in the exploration part of the full interaction state, which, we expect, could affect the intuitiveness of the interaction. However, in a preliminary experiment, the participants rated the speed of the robot as good. An interesting research question could be what is the proper movement speed a robot might exhibit in order to be perceived as harmless. We included this topic in the future development of the experiment. Furthermore, in the current experiment, although the fastest processing was in interaction avoidance, people perceived the robot as less responsive than during interaction and full interaction.

3.4 Participants

In total 28 people participated in the survey, which results in a total of 112 questionnaires (four questionnaires per participant, one for each interaction). Some participants missed to answer some questions, but those were only a few questions. It is interesting to note that few participants had negative or neutral responses in all four experiments, regardless of the experiment, together with comments saying that Nao did not want anything because it is a machine. This might be perceived as a negative bias towards robots.

Of 28 participants, 8 were female (28.57 %) and 20 were male (71.43 %). There were 17 Germans, 2 Italians, 2 Serbians, 2 Poles, 1 Czech, 1 Dutch, 1 Estonian and 1 French. Regarding previous experience with robots, 25 persons (89.29 %) had none and 3 (10.71 %) had previous experience—one with industrial robots, one with Aldebaran Nao and one with Lego Mindstorms. The average age of the participants was 28.12 ($\sigma = 5.64$). Among the participants, 75 % had university level education and 25 % had highschool level education.

Unfortunately, not all the participants allowed to be filmed during their interaction because of privacy reasons (even though we informed them that the data will be kept anonymous and videos will not be published against their wish). The video database is composed of 10 videos for *exploration*, 7 for *interaction*, 8 for *avoid interaction* and 9 for *full interaction*.

3.5 Measurements

Only recently, performance criteria different from those typical for industrial robots have been adopted for measuring the success of social and service robots. Current criteria lie within the satisfaction of the user [18].

We decided to adopt two techniques for evaluating the interaction: questionnaires and proxemics estimated from recorded video sequences of the interaction. So far, we wanted to adopt only metrics related to socio-cognitive skill perception instead of measuring the affective state of the user through the use of physiological sensors.

3.5.1 Questionnaires

We conducted a qualitative, anonymous survey to evaluate how people perceive their interaction with the Nao. Questionnaires are often used to measure the user's attitude. The first encountered problem was related to what type of questionnaire to adopt. Developing a valid questionnaire can take a considerable amount of time and the absence of standardisation makes it difficult to compare the results with other studies. That is why we decided to adopt standardised measurement tools for human-robot interaction, in addition to some metrics we found interesting for our research. We adopted as a part of our survey the Godspeed questionnaire [7] which uses semantic differential scales for evaluating the attitude towards the robot. Such a questionnaire contains questions (variables) about five concepts (latent variables): Anthropomorphism, Animacy, Likeability, Perceived Intelligence and Perceived Safety (for a detailed description and for the set of questions, please refer to [7]).

Anthropomorphism refers to the attribution of human features and behaviours to non-human agents, such as animals, computers or robots. Anthropomorphism variables were (left value scored as 1, right value scored as 5): fake–natural, machinelike–humanlike, unconscious–conscious, artificial–lifelike, moving rigidly–moving elegantly.

Animacy is the property of alive agents. Robots can perform physical behaviours and reactions to stimuli. The participants' perception about robot animacy can give important insights for improving robot skills. Variables were: dead–alive, stagnant–lively, mechanical–organic, artificial–lifelike (different from the one in anthropomorphism, as related to the animacy), inert–interactive, apathetic–responsive.

Likeability may influence the user's judgments. Some studies indicate that people often make important judgments within seconds of meeting a person and it is assumed that people are able to judge also a robot [7]. Likeability variables were: dislike–like, unfriendly–friendly, unkind–kind, unpleasant–pleasant, awful–nice.

Perceived Intelligence is one of the most important metrics for evaluating the efficacy of the implemented skills. It can depend on robot competence, but the duration of the interaction is also one of the most influencing factors, as users can become bored if the interaction is long and the vocabulary of the robot's behaviours is limited. Variables were: incompetent–competent, ignorant–knowledgeable, irresponsible–responsible, unintelligent–intelligent, foolish–sensible.

Perceived Safety is a metric for estimating the user's level of comfort when interacting with the robot and the perception of the level of danger. Variables were: anxious–relaxed, agitated–calm, quiescent–surprised (this variable was recoded, as explained in the next paragraph).

The reliability of the questionnaire was analysed by its authors, who claim that such questions have sufficient internal consistency and reliability; to confirm this, we computed Cronbach's alpha[1] for each latent variable again. We found that Cronbach's alpha was negative ($\alpha = -1.111$) for the latent variable Perceived Safety, due to a negative average covariance among items. This violated reliability model assumptions for that set of variables, due to a miscoding of a variable. In fact, the questionnaire is written in such a way that high values of one variable mean the same thing as low values of the other variable; the miscoded variable was: Quiescent (scaled as 1) to Surprised (scaled as 5), probably due to the fact that participants intended quiescence as a synonym for calmness (the previous variable was Agitated, coded as 1, or Calm, coded as 5). After recoding the quiescent–surprised variable, the Cronbach's alpha proved to be higher ($\alpha_{Perceived\ Safety} = 0.839$).[2] We did not find any other problems with the rest of the latent variables: $\alpha_{Anthropomorphism} = 0.825$, $\alpha_{Animacy} = 0.853$, $\alpha_{Likeability} = 0.813$, $\alpha_{Perceived\ Intelligence} = 0.750$.

In addition to the Godspeed questionnaire, we introduced a new latent variable for measuring the concept of User Satisfaction, with two variables: frustrating—exciting and unsatisfying interaction—satisfying interaction (high Cronbach's alpha: $\alpha_{User\ Satisfaction} = 0.799$).

Open questions were also introduced about the understanding of the behaviour of the robot, its desires, its aiming to interact or not, its successfulness, its gender (with the

---

[1]High Cronbach's alpha values are those greater than 0.5, which specify that the used set of variables are good for defining a certain concept.

[2]Recoding represents inversion of the variable in the following manner: $1 = 5$, $2 = 4$, $3 = 3$, $4 = 2$, $5 = 1$.

explanation of the chosen one), its age, type of communication during the interaction, expectations about future improvements and differences between Nao and humans.

3.5.2 Proxemics

According to the sociological concept of proxemics, humans, as well as animals, use to define personal spheres which delimit areas of physical distance that correlate reliably with how much people have in common [19]. The boundaries of such spheres are determined by factors like gender, age and culture. Coming inside the sphere of another person may let him/her feel intimidated, or staying too far can be seen as cold or distant. Four spheres were identified, according to [19]: Intimate Distance (from 0 to 45 cm), reserved for embracing, touching, whispering; Personal Distance (from 45 to 120 cm), reserved for friends; Social Distance (from 1.2 to 3.6 m), reserved for acquaintances and strangers; Public Distance (more than 3.6 m), reserved for public speaking.

To the best of our knowledge, in human-robot interaction, no assumptions about the existence of such boundaries have been made. The focus has been pointed on identifying those factors that influence interaction distance. Interaction distance can be influenced by factors like user age or gender, pet ownership, crowdedness in the environment or available space, as shown in [8, 19]. However, their analyses did not include users' perceptions about the behaviour or features of the robot.

We wanted to include proxemics measurement hoping to find some correlations between interaction distance and the factors treated in the questionnaire. We analysed participant behaviour also from measuring the distances between the face of the robot and the face of the user and between the face of the robot and the hand of the user.[3]

As introduced in Sect. 3.2, proxemics analyses were done by gathering data from video recorded during the interaction sessions (Fig. 2 shows a sample frame). The user sat in front of the robot at a distance of ca. 90 cm. We recorded the interaction with a standard camera (resolution 640×480) placed at ca. 2 meters perpendicularly to the axis robot-user. Beside the table where the robot was standing there was a scale drawn on a whiteboard for the visual estimation (estimated average error: 5 cm) of the distance between the nose of the user and the head of the robot and from the hand of the user and the head of the robot. Videos were annotated manually: every 5 seconds the face-face and face-hand distances were visually estimated by the operator, manually projecting their positions onto the scale drawn on the whiteboard.

Participants were sitting on a chair (they all started at the same distance to the robot), but they were told to feel free to interact in any way they considered more appropriate. However, it happened only in very few cases (only 2 participants) that they stood up. In both cases, we gathered the face-face and face-hand distances as projected onto the horizontal line parallel to the table.

3.6 Results

This section presents the quantitative evaluation of our experiments.

In an earlier experiment we noticed some interesting patterns [4, 9]. It seemed that if a person holds the object close to the robot's hand, then Nao's pointing will be perceived as a desire to grasp the object. This could indicate, along with the hypothesis that pointing emerges from grasping, that there is also a reverse connection—pointing can be perceived as grasping, if the object is too close to the hand.[4] Furthermore, most of the participants in the preliminary experiment responded that Nao was either likeable or very likeable and that the speed of experiment was good (out of three possible answers: too fast, good and too slow), even though the execution speed was lower than in the current experiment. All participants in the preliminary experiment, except one, had no previous experience with robots.

Figure 3 shows the means and the standard deviations of the responses.

First, we checked whether the distributions of the collected data are normal or not, in order to select the proper statistical tests. For each variable (that is, for each question), we looked at the superimposition of the histogram of the data with a normal curve characterised by the mean and the variance of the data. Almost all the histograms did not fit well together with the corresponding normal curves. Thus, we checked the kurtosis and the skewness of the data,[5] in order to have a more precise measurement of the normality of the distributions. The distributions of all the variables related to the questionnaire had kurtosis and skewness between -2 and $+2$, while 17 out of 64 distributions related to the variables of the proxemics analysis[6] did not.

---

[3]When interacting with the robot, participants did not use two hands at the same time. Almost all of them performed movements only with one arm, or at least they alternated between left and right. We registered only the movements from the active one.

[4]The robot platform we used has no movable fingers and it is unable to grasp an object.

[5]In general, when kurtosis and skewness are between -2 and $+2$, the data is not too far away from a normal distribution. When that is not the case, corrections (like Box-Cox transformations) can be applied to the data in order to apply the tests that have assumptions of normality.

[6]For each of the four behaviours performed by the robot, we created two variables for the average value and variance of the distance between the face of the Nao and the nose of the participant for the following cases: during the first 15 seconds of the interaction, between the 15th second and the 45th second of the interaction, and during the last 15 seconds of the interaction (in total 6 variables). The same variables were created for analysing the distance between the face of the Nao and the user's hand closest to the robot.
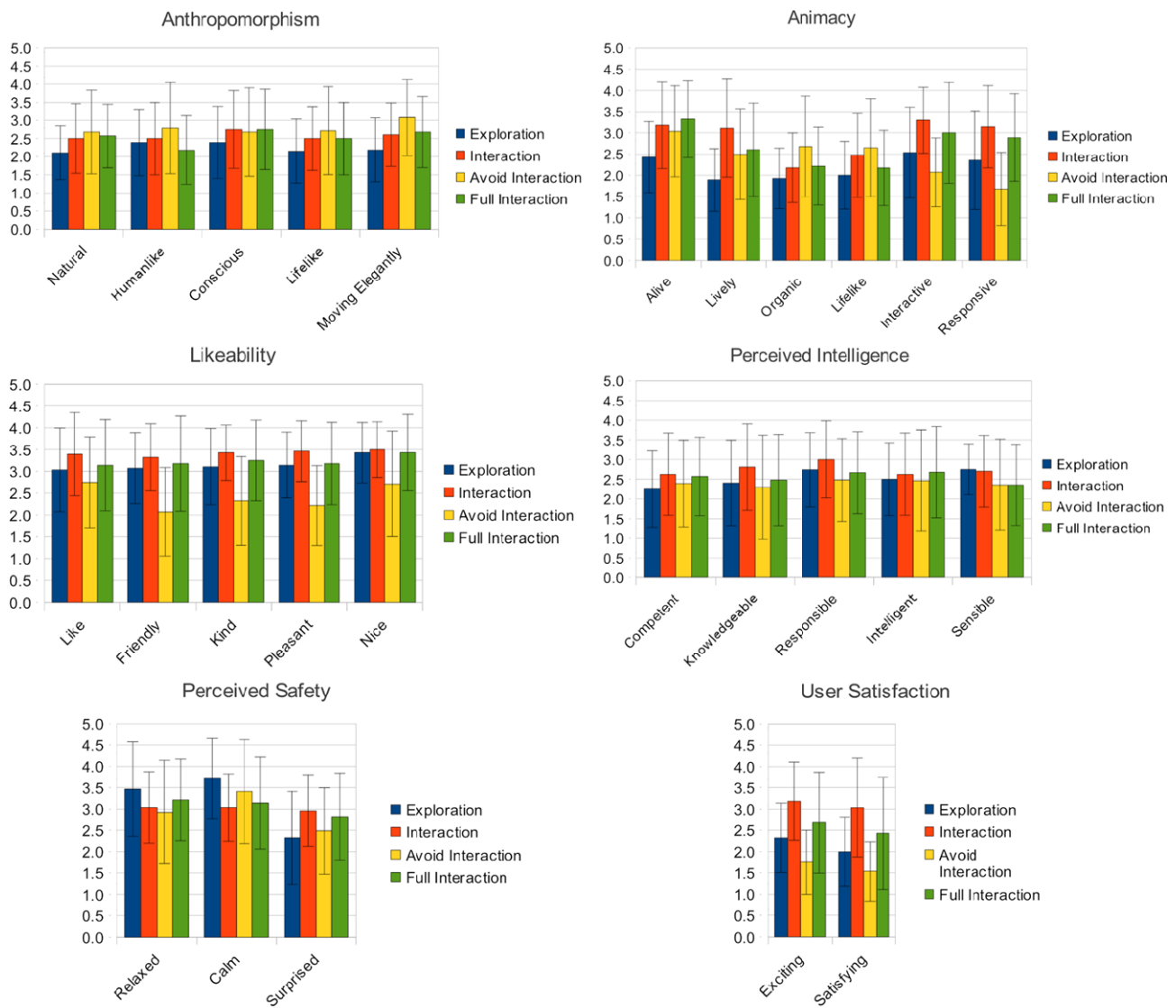
Fig. 3 These graphs show the results taken from the Godspeed questionnaire

Due to the non-normality of such distributions, it seems to be more appropriate applying non-parametric statistical tests for the whole analysis. However, the use of ANOVA on Likert-scale data and without the assumption of normality of the distributions of the data to be analysed is controversial. In general, researchers claim that only non-parametric statistics should be used on Likert-scale data and when the normality assumption is violated. Vallejo et al. [20], instead, found that the Repeated Measures ANOVA[7] was robust toward the violation of normality assumption. Simulation results of Schmider et al. [21] confirm also this observation, since they found in their Monte Carlo study that the empiri-

cal Types I and II errors in ANOVA were not affected by the violation of assumptions.

3.6.1 Correlations

A Spearman's Rank Order correlation[8] was run to determine the relationship between perceived factors and between them and average human-robot distances. Each run was done for each experimental session (*exploration*, *interaction*, *interaction avoidance* and *full interaction*).

---

[7]Repeated measures ANOVA compare the average score for a single group of subjects at multiple time periods (observations).

[8]Spearman's correlation coefficient is non-parametric, looks at ranked (coded) variables (without looking at the data directly) and does not have the normality assumption on the distributions, thus it can be used for skewed or ordinal variables. We ran the correlation with 2-tailed test of significance. Missing values were excluded with cases pairwise.

Tables 1 and 2 show some of the most relevant correlations. In addition to the data shown in the tables, it has to be noted that in the *exploration* test there was a strong, positive correlation between almost all the anthropomorphism variables and the perceived intelligence attributes related to competence and knowledge; in *interaction*, the higher the likeability of the robot, the higher the variance of face-face distance during all the interaction tests ($r = 0.805$, $P = 0.029$, $N = 7$); in *full interaction*, perceived intelligence was found to be positively correlated with almost all the other variables (except those related to perceived safety) with $r > 0.5$ and almost always significant at the 0.01 level.

3.6.2 Repeated Measures ANOVA

Because the participants of the four different observations were the same in each group, we adopted the Repeated measures ANOVA test (post-hoc test using Bonferroni correction) for the analysis of variances. Also known as within-subjects ANOVA test, repeated measures ANOVA is the equivalent of the one-way ANOVA but for related, not independent groups. We performed the test on all the dependent variables.[9]

Post-hoc tests revealed that the four different behaviours performed by the robot have not changed significantly the participants' perception of the anthropomorphic attributes related to naturalness, humanlikeness, consciousness and artificiality. Table 3 shows the statistically significant results of repeated measures ANOVA on the questionnaire variables.

Proxemics variables contain a high number of missing values. In order to perform repeated measures ANOVA on those variables, we had to replace missing values with multiple imputation ($n = 20$). New samples were created, where proxemics information was inferred using the questionnaire variables as predictors.[10]

Table 4 shows the statistically significant results of the repeated measures ANOVA on the proxemics variables.

3.6.3 Latent Growth Curve Model

A latent growth curve model was also used to assess the change in user perception over the four behaviours. This model uses a structural equation to estimate two latent variables, the slope and intercept, to assess the average linear change across the measurements, where the individual measurements are the indicators of the latents.[11] The estimated population distribution of the linear change (or growth) trajectory, denoted by the slope and the intercept of a linear function, are derived from this structural equation model. The estimator selected for the procedure was a Bayesian estimator with non-informative priors.[12] All calculations were produced with Mplus 6.11.

The estimated slopes for many of the items were almost all positive, with also positive credibility intervals, meaning that there is a significant positive trend in the average score from the first observation (*exploration*) to the last observation (*full interaction*).[13]

3.7 Discussion

Despite the small sample size of the data collected during the experiments (especially regarding the proxemics analysis), the outcomes suggested many elements and features that need to be carefully taken into account when developing attentive mechanisms for intuitive robot behaviour.

3.7.1 Godspeed questionnaire

The adoption of the Godspeed questionnaire allowed us to test its qualities. Questionnaires are important tools in measuring user perceptions and the Godspeed provided us with a good instrument for measuring the quality of the implemented robot behaviours. Its authors noted that comparing different robots and their settings by means of the same measurement index will help roboticists in making design decisions. In [22], the indices of the Godspeed questionnaire have been tested as measures of human-like characters. The

---

[9]Mauchly's test has been used as statistical test for validating repeated measures ANOVA. It tests the sphericity, which is related to the equality of the variances of the differences between levels of the repeated measures factor. Sphericity, an assumption of repeated measures ANOVA, requires that the variances for each set of difference scores are equal. Sphericity can not be assumed when the significance level of the Mauchly's test is < 0.05. Violations of sphericity assumption can invalidate the analysis conclusions, but corrections can be applied to alter the degrees of freedom in order to produce a more accurate significance value, like the Greenhouse-Geisser correction. When the significance level of the Greenhouse-Geisser estimate is < 0.05, statistical significant differences revealed by post-hoc test can be elicited from the pairwise comparisons between the observations. Repeated Measures ANOVA does not tell where the differences between groups lie. When repeated measures ANOVA is statistically significant (both with sphericity assumption not violated or with Greenhouse-Geisser correction), post-hoc tests with multiple comparisons can highlight exactly where these differences occur.

[10]For multiple imputation, all the available variables that can predict the values of missing data should be included.

[11]The loadings are constrained to be 1 for the intercept latent and to 0 to 3 (depending of the time of measurement) for the slope latent.

[12]This estimation strategy was appropriate as the more commonly used maximum likelihood estimator often produces biased (or often inestimable) results with such small sample sizes. The Bayesian estimator is more robust to both small samples and violation of distributional assumptions that could emerge from small samples.

[13]Further analysis can be done on piecewise linear growth, for breaking up the curvilinear growth trajectories into separate linear components, thus for analysing whether there was an increase or a decrease between *exploration* and *interaction*, between *interaction* to *interaction avoidance*, and so on.

Table 1 Most relevant correlations (Part 1). For having the full tables, please ask the authors

Variables correlated		Exploration			Interaction			Inter. avoidance			Full interaction		
		R	p	N	R	p	N	R	p	N	R	p	N
Anthropomorphism: humanlike	Animacy: alive	0.581	0.001	28	0.654	0.000	28	0.665	0.000	28	0.546	0.003	28
Anthropomorphism: humanlike	Animacy: interactive	0.513	0.005	28	0.605	0.001	27	Stat. not signif.			0.504	0.006	28
Anthropomorphism: humanlike	Perc. Intelligence: knowledgeable	0.416	0.031	28	0.562	0.003	26	0.571	0.002	27	0.606	0.001	27
Anthropomorphism: humanlike	Perceived Intelligence: competent	0.476	0.011	28	0.677	0.000	27	0.623	0.000	28	0.564	0.002	28
Anthropomorphism: humanlike	Perceived Intelligence: intelligent	Stat. not signif.			0.559	0.002	27	0.573	0.001	28	**0.713**	0.000	28
Anthropomorphism: natural	Perc. Intelligence: knowledgeable	0.498	0.008	27	0.557	0.003	26	Stat. not signif.			0.553	0.003	27
Anthropomorphism: natural	Perc. Intelligence: competent	0.565	0.002	28	0.612	0.001	27	Stat. not signif.			0.572	0.001	28
Anthropomorphism: moving elegantly	Perc. Intelligence: knowledgeable	0.697	0.000	26	0.399	0.044	26	0.654	0.000	27	0.422	0.028	27
Anthropomorphism: moving elegantly	Perceived Intelligence: competent	**0.694**	0.000	27	0.483	0.011	27	0.542	0.003	28	0.454	0.015	28
Anthropomorphism: moving elegantly	Likeability: friendly	Stat. not signif.			Stat. not signif.			−0.500	0.007	28	Stat. not signif.		
Anthropomorphism: lifelike	Variance Face-Hand (15"–45")	Stat. not signif.			Stat. not signif.			Stat. not signif.			**−0.786**	0.012	9
Animacy: lifelike	Likeability: friendly	Stat. not signif.			0.663	0.001	23	−0.425	0.043	23	Stat. not signif.		
Animacy: interactive	Anthropomorphism: lifelike	**0.673**	0.012	13	0.660	0.000	27	Stat. not signif.			0.556	0.002	28
Animacy: interactive	Likeability: friendly	0.398	0.036	28	0.451	0.018	27	Stat. not signif.			0.655	0.000	28
Animacy: interactive	Perceived Intelligence: intelligent	**0.462**	0.013	28	**0.705**	0.000	26	**0.403**	0.033	28	**0.619**	0.000	28
Animacy: interactive	User Satisf.: exciting	**0.710**	0.000	28	0.551	0.004	26	Stat. not signif.			**0.706**	0.000	28
Animacy: interactive	User Satisf.: satisfying	0.470	0.012	28	0.687	0.000	26	Stat. not signif.			0.725	0.000	28
Animacy: responsive	Average Face-Face distance (60 s.)	0.633	0.037	11	Stat. not signif.			Stat. not signif.			Stat. not signif.		
User Satisf.: satisfying interaction	Anthropomorphism: moving elegantly	0.505	0.007	27	0.482	0.011	27	Stat. not signif.			0.390	0.040	28
User Satisf.: satisfying interaction	Anthropomorphism: lifelike	0.576	0.002	26	**0.653**	0.000	27	Stat. not signif.			0.516	0.005	28
User Satisf.: satisfying interaction	Animacy: responsive	**0.696**	0.000	28	**0.722**	0.000	27	Stat. not signif.			**0.673**	0.000	28
Perceived Safety: quiescent	Average Face-Face dist. (last 15")	Stat. not signif.			Stat. not signif.			−0.879	0.009	7	Stat. not signif.		
Perceived Safety: quiescent	Average Face-Hand dist. (last 15")	Stat. not signif.			Stat. not signif.			−0.805	0.029	7	Stat. not signif.		

results indicated significant and strong correlations among some relevant indices and new indices have been proposed. This matches the comments of most of the participants of our experiments which complained about the similarity between many questions and about some high-level attributes which were difficult to assign to the robot. The problem we

reported with the recorded variable and the previous notes suggest to not adopt the original version of the Godspeed questionnaire for further experiments, but rather its revisited version.

To the best of our knowledge, no other study on attentional mechanisms for robots has adopted the Godspeed

Table 2 Most relevant correlations (Part 2). For having the full tables, please ask the authors

Variables correlated		Exploration			Interaction			Inter. avoidance			Full interaction		
		R	p	N	R	p	N	R	p	N	R	p	N
Variance Face-Hand dist. (0"–15")	Perceived Safety: quiescent	Stat. not signif.			Stat. not signif.			Stat. not signif.			−0.670	0.048	9
Variance Face-Hand dist. (60 s.)	Likeability: friendly	Stat. not signif.			0.802	0.030	7	Stat. not signif.			−0.673	0.047	9
Variance Face-Hand dist. (60 s.)	Likeability: kind	Stat. not signif.			Stat. not signif.			Stat. not signif.			**−0.738**	0.023	9
Variance Face-Hand dist. (60 s.)	Likeability: pleasant	Stat. not signif.			Stat. not signif.			Stat. not signif.			**−0.829**	0.006	9
Variance Face-Hand dist. (60 s.)	User Satisf.: satisfying interaction	Stat. not signif.			Stat. not signif.			Stat. not signif.			**−0.738**	0.023	9
Variance Face-Face dist. (15"–45")	Likeability: friendly	Stat. not signif.			0.809	0.028	7	Stat. not signif.			Stat. not signif.		
Average Face-Face dist. (60 s.)	User Satisf.: exciting	Stat. not signif.			Stat. not signif.			Stat. not signif.			0.709	0.032	9
Average Face-Face dist. (60 s.)	Perc. Intellingence: intelligent	Stat. not signif.			Stat. not signif.			Stat. not signif.			0.729	0.026	9

questionnaire as a metric. However, in [23], the authors studied the combined and individual contribution of gestures and gazing to the persuasiveness of a story-telling robot measuring user perception with the Godspeed questionnaire. The robots used persuasive gestures (or not) to accompany the persuasive story, and also used gazing (or not) while telling the persuasive story. Their results indicated that only gazing had a main effect on persuasiveness, while the use of gestures did not. Also, the combined effect of gestures and gazing on persuasiveness was greater than the effect of either gestures or gazing alone. This study suggests that adding multiple social cues can have additive persuasive effects, matching what we will discuss in the next subsection about multi-modal interaction and efficient feedback systems.

3.7.2 Correlations

Correlation analysis confirmed our expectations and suggested directions for improvement of robot attention mechanisms. Positive correlations between anthropomorphic attributes and perceived intelligence confirmed that a robot with human-like appearance can increase the level of its perceived intelligence. However, an excessive human-like appearance can entail the interacting person having too high expectations about the robot's cognitive capabilities, which can provoke disappointment whenever the robot does not fulfill such expectations. We believe that the positive correlations between the anthropomorphic attributes and the perceived intelligence reflect a good balancing between Nao's human-like appearance and its implemented cognitive capabilities. Confirming this hypothesis, most of the

participants did not try to communicate vocally with the robot, suggesting that they were not expecting this interaction modality due to the absence of a mouth in the robot's face and due to any other robot's verbal capability.

Positive correlations between the robot's interactiveness and user excitement and perception of lifelikeness and intelligence (see Table 1, correlations between Animacy: interactive and Perceived Intelligence: intelligent) suggested also that interactive capabilities emerging from attention mechanisms can increase the perceived level of intelligence of the robot. Such results confirm also the thought that a robot has to be highly interactive for being perceived as a highly intelligent agent, and it has to be responsive for increasing user satisfaction.

We believe that a relevant contribution in the user satisfaction is given by the robot's responsiveness and interactivity and it can be increased by improving its feedback system. A well designed feedback system could reduce the consequences of some of the robot limitations. In our experiments, participants experienced issues related to the limited field of view of the Nao (58° diagonal FOV). It is plausible that humans expect of humanoids to have approximately matching characteristics, such as the field of view, or two eyes for vision.[14] During the experiments, participants, without being aware of that, were often waving to the robot or handing over the object out of the robot's field of view causing no reactions to it. This resulted in affecting the perception of the robot's responsiveness and inter-

---

[14] Video cameras are not located in the positions of the eyes on the Nao, which leads to unmet expectations.

Table 3 Statistically significant results of repeated measures ANOVA on the questionnaire variables. Cases with sphericity assumption violated were corrected with Greenhouse-Geisser method. The *table* shows the statistically significant pairwise comparisons (illustrating the changes in means from an observation to another), taken from the post-hoc test with Bonferroni correction

Variable	Sphericity assumed	From observ.	To observ.	Mean difference	Std. error	Significance
Anthr.: moving elegantly	no	1	3	−0.889	0.252	0.010
		2	3	−0.481	0.154	0.026
Animacy: alive	yes	1	2	−0.75	0.203	0.006
		1	4	−0.893	0.165	0.000
Animacy: lively	yes	1	2	−1.222	0.284	0.001
		1	4	−0.741	0.224	0.016
Animacy: organic	no	1	2	−0.750	0.239	0.025
		2	3	−0.500	0.159	0.024
Animacy: interactive	yes	1	2	−0.815	0.251	0.019
		2	3	1.222	0.202	0.000
		3	4	−1.000	0.233	0.001
Animacy: responsive	yes	1	2	−0.786	0.259	0.032
		2	3	1.464	0.260	0.000
		3	4	−1.214	0.243	0.000
Likeability: friendly	no	1	3	1.000	0.230	0.001
		2	3	1.250	0.270	0.001
		3	4	−1.107	0.274	0.002
Likeability: kind	yes	1	3	0.786	0.243	0.019
		2	3	1.107	0.248	0.001
		3	4	−0.929	0.224	0.002
Likeability: pleasant	no	1	3	0.929	0.185	0.000
		2	3	1.250	0.222	0.000
		3	4	−0.964	0.238	0.002
Likeability: nice	no	1	3	0.714	0.198	0.008
		2	3	0.786	0.249	0.023
Perceived Safety: quiescent	yes	1	2	0.593	0.194	0.031
		2	3	−0.481	0.154	0.026
User satisfaction: exciting	no	1	2	−0.852	0.218	0.004
		2	3	1.407	0.234	0.000
		2	4	0.444	0.154	0.047
		3	4	−0.963	0.285	0.014
User satisfaction: satisfying	yes	1	2	−1.037	0.196	0.000
		2	3	1.519	0.222	0.000
		3	4	−0.963	0.229	0.002

activeness. A little foresight in the feedback system could have probably reduced this effect, like changing the colour of head LEDs, or emitting sounds, whenever the robot detected something.

Multi-modal interaction (through arm or head movements) increased the level of interactiveness perceived by participants, as suggested by the correlations between Animacy:interactive and several other variables (see Table 1)

Table 4 Statistically significant results of repeated measures ANOVA on the proxemics variables. Cases with sphericity assumption violated were corrected with Greenhouse-Geisser method. The *table* shows the statistically significant pairwise comparisons (illustrating the changes in means from an observation to another), taken from the post-hoc test with Bonferroni correction. Missing values were replaced with multiple imputations. The new dataset contained 560 samples. Abbreviations: AV: average; VAR: variance; FF: distance between the face of the robot and the face of the user; FH: distance between the face of the robot and the closest hand of the user; all: considering the whole duration of the test (60 seconds)

Variable	Sphericity assumed	From observ.	To observ.	Mean difference	Std. error	Significance
AV FF all	no	1	2	**13.675**	0.419	0.000
		1	3	11.876	0.302	0.000
		1	4	9.734	0.355	0.000
		2	3	**−1.799**	0.360	0.000
		2	4	**−3.941**	0.436	0.000
		3	4	−2.142	0.280	0.000
VAR FF all	no	1	2	−6.218	2.058	0.016
		1	3	−82.552	3.014	0.000
		1	4	**14.558**	1.914	0.000
		2	3	−76.334	2.361	0.000
		2	4	20.776	1.633	0.000
		3	4	97.110	2.862	0.000
AV FH all	no	1	2	14.767	0.544	0.000
		1	3	**18.439**	0.547	0.000
		1	4	**21.423**	0.539	0.000
		2	3	3.672	0.294	0.000
		2	4	6.656	0.314	0.000
		3	4	2.985	0.307	0.000
VAR FH all	no	1	2	**20.337**	3.861	0.000
		1	3	**−217.131**	7.246	0.000
		2	3	**−237.468**	6.718	0.000
		2	4	−28.076	5.602	0.000
		3	4	**209.392**	6.904	0.000

which during *interaction* were higher than when the robot performed other behaviours.[15] The consideration of [23] about combining gestures and gazing for increasing the persuasiveness and the likeability of the robot matches our consideration about multi-modal interaction. Elegance in movement positively correlating with user satisfaction suggests that the robot might perform smooth and natural movements in order to increase the quality of the interaction.

A trustworthy and lifelike robot can be better accepted as a companion or as a co-worker, where close interaction is needed, as suggested by the negative correlation between lifelikeness and face-face average distance recorded during *interaction* ($r = 0.805$, $P = 0.029$, $N = 7$).

---

[15]During *interaction*, the robot performed arm and head movements during the whole session.

3.7.3 Repeated Measures ANOVA

Repeated measures ANOVA results showed that the aliveness of the robot during *exploration* scored lower than during *interaction* and *full interaction*, again supporting our expectation that multi-modal interaction increases the expressiveness of the robot behaviours (in *exploration*, the robot performed only head movements). Again, more expressive movements or a better designed feedback system could have increased the level of perceived animacy, likeability and user satisfaction.

The less the interaction was perceived as satisfactory, the more often and the more frenetically the participants moved their hand. Repeated measures ANOVA confirmed that the variance of face-hand distance is higher during *interaction avoidance* (the least satisfactory robot behaviour for the users) than during the other behaviours. It is also interesting to note how successful the *interaction avoidance* behaviour was, by which the robot did cause frustration to

the users, according to its *motivation* of avoiding the interaction. Several participants commented this behaviour assigning mental states to the robot, like *shyness* and *angriness*.

4 Conclusions

We have created a saliency based attentional model combined with a robot ego-sphere and implemented it on a humanoid robot. In human-robot interaction experiments using this model, we show that different attentional behaviour of the robot has a strong influence on the interaction as experienced by the human.

We have shown that—even on robots with limited computational capacities such as the Nao—it is possible to have an ongoing interaction between the robot and a person. Techniques used are a combination of bottom-up and top-down processes of attention and an ego-sphere as a short-term memory representation in combination with motion, face and object detection.

The adopted questionnaires were useful for correlating perceived physical and behavioural robot features with proxemics data. We noticed some trends suggesting that some of the perceived variables could influence the distances of the interaction.

Through the discussion of the results in the previous section, we identified those characteristics that need to be emphasised and those skills that have to be taken into account (like providing enough feedback during the interaction) when implementing attentive mechanisms in robots.

For future experiments we plan to explore different approaches for dynamic weight assignment for different filters. We also plan to extend the system to include more different filters on the Nao robot (e.g. for audio localisation), as well as to port the approach to other robot platforms. It would be interesting to see how these attentional models would rate on other, non-humanoid platforms. Additionally, the presented *full interaction* behaviour, consisting of *exploration*, *interaction* and *interaction avoidance*, can be applied to more complex scenarios, and we are planning to explore this further. Gesture recognition and synthesis, and behaviour recognition and execution would enable the robot to better communicate its intentions and understand the intentions of others. We believe that giving visual and auditory feedback to the participant is of extreme importance for increasing the intuitiveness of the interaction and the user satisfaction.

Another interesting research question could be what is the proper movement speed a robot might exhibit in order to be perceived harmless. We included this topic in the future development of the experiment.

We believe that these experiments represent a step in a good direction toward reaching joint attention between a human and a robot. We showed that basic attention manipulation is possible, even with simple robot platforms, such as the Nao, and that participants will assign different characteristics to it based on its behaviour.

References

1. Burghart CR, Steinfeld A (eds) (2008) Proceedings of the metrics for human-robot interaction workshop at the 3rd ACM/IEEE international conference on human-robot interaction (HRI 2008)
2. Kaplan F, Hafner VV (2006) The challenges of joint attention. Interact Stud 7(2):135–169
3. Schillaci G, Hafner VV (2011) Prerequisites for intuitive interaction—on the example of humanoid motor babbling. In: Proceedings of the workshop on the role of expectations in intuitive human-robot interaction (HRI 2011), pp 23–27
4. Bodiroza S, Schillaci G, Hafner VV (2011) Robot ego-sphere: an approach for saliency detection and attention manipulation in humanoid robots for intuitive interaction. In: Proceedings of the 11th IEEE-RAS conference on humanoid robots, pp 689–694
5. Jonsson GK, Thorisson KR (2010) Evaluating multimodal human-robot interaction: A case study of an early humanoid prototype. In: Proceedings of the 7th international conference on methods and techniques in behavioral research, Ser. MB '10. ACM, New York, pp 9:1–9:4
6. Steinfeld A, Fong T, Kaber D, Lewis M, Scholtz J, Schultz A, Goodrich M (2006) Common metrics for human-robot interaction. In: Proceedings of the 1st ACM SIGCHI/SIGART conference on human-robot interaction, Ser. HRI '06. ACM, New York, pp 33–40
7. Bartneck C, Kulić 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(1):71–81
8. Takayama L, Pantofaru C (2009) Influences on proxemic behaviors in human-robot interaction. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp 5495–5502
9. Hafner VV, Schillaci G (2011) From field of view to field of reach—could pointing emerge from the development of grasping? In: Proceedings of the IEEE conference on development and learning and epigenetic robotics (IEEE ICDL-EPIROB 2011), conference abstract in frontiers in computational neuroscience
10. Tomasello M (1995) Joint attention as social cognition. In: Moore C, Dunham PJ (eds) Joint attention: its origins and role in development, vol 16, pp 103–130, Erlbaum
11. Itti L, Koch C (2001) Computational modelling of visual attention. Nat Rev Neurosci 2:194–203
12. Treisman A (1985) Preattentive processing in vision. Comput Vis Graph Image Process 31(2):156–177
13. Peters RA, Hambuchen KE, Kawamura K, Wilkes DM (2001) The sensory ego-sphere as a short-term memory for humanoids. In: Proceedings of the IEEE-RAS conference on humanoid robots, pp 451–460

14. Fleming KA, Peters RA, Bodenheimer RE (2006) Image mapping and visual attention on a sensory ego-sphere. In: IEEE/RSJ international conference on intelligent robots and systems, pp 241–246

15. Ruesch J, Lopes M, Bernardino A, Hornstein J, Santos-Victor J, Pfeifer R (2008) Multimodal saliency-based bottom-up attention a framework for the humanoid robot iCub. In: IEEE international conference on robotics and automation. ICRA 2008, pp 962–967

16. Schillaci G, Hafner VV (2011) Random movement strategies in self-exploration for a humanoid robot. In: Proceedings of the 6th ACM/IEEE international conference on human-robot interaction (HRI 2011), pp 245–246

17. Burkhard H, Holzhauer F, Krause T, Mellmann H, Ritter C, Welter O, Xu Y (2010) NAO-Team Humboldt, Humboldt-Universität zu Berlin

18. Bartneck C, Croft E, Kulic D (2008) Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots. University of Hertfordshire, pp 37–44

19. van Oosterhout T, Visser A (2008) A visual method for robot proxemics measurements. In: Proceedings of the metrics for human-robot interaction workshop in affiliation with the 3rd ACM/IEEE international conference on human-robot interaction (HRI 2008). Technical Report 471, pp 61–68

20. Vallejo G, Fernández MP, Tuero E, Livacic-Rojas P (2010) Análisis de medidas repetidas usando métodos de remuestreo. Anal Psicol 26(2):400–409. Engl translate: Analyzing repeated measures using resampling methods

21. Schmider E, Ziegler M, Danay E, Beyer L, Bühner M (2010) Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption. Methodology 6(4):147–151

22. Ho C, MacDorman K (2010) Revisiting the uncanny valley theory: developing and validating an alternative to the godspeed indices. In: Computers in Human Behavior, vol 26, pp 1508–1518

23. Ham J, Bokhorst R, Cabibihan J (2011) The influence of gazing and gestures of a storytelling robot on its persuasive power. In: International conference on social robotics

Guido Schillaci received his B.Sc. (2004) and M.Sc. (2009) degrees in Computer Engineering from the University of Palermo, Italy. In 2007, he studied as an exchange student at ETSIIT, University of Granada, Spain. In 2009, he was granted a one-year research scholarship by the Italian Government inherent to the project FRamework for Agent-based Semantic-aware Interoperability. Since 2010, he is a Marie Curie fellow supervised by Prof. Dr. V.V. Hafner at Humboldt-Universität zu Berlin, Germany. He

is associate fellow of the Marie Curie ITN RobotDoC, IEEE Student Member and EUCog Member. Embodied cognition, sensorimotor learning and internal simulations, developmental robotics, social interaction and theory of mind are milestones for his current research.

Saša Bodiroža received his B.Sc. (2009) and M.Sc. (2010) degrees in Computer Science and Engineering from the School of Electrical Engineering, University of Belgrade, Serbia. In 2010 he started his doctoral work with Prof. Dr. Verena Hafner as a Marie Curie fellow and a member of the Cognitive Robotics Group at Humboldt-Universität zu Berlin, Germany. His research interests are in the broad field of HRI and particularly the use of gestures in HRI, including dynamic gesture recognition, learning and synthesis, as well as optimal gesture selection.

Verena Vanessa Hafner is Junior Professor at Humboldt-Universität zu Berlin and head of the Cognitive Robotics Group at the Department of Computer Science since 2007. She holds a Masters in Computer Science and AI with distinction from the University of Sussex (UK), and a PhD from the Artificial Intelligence Lab, University of Zurich (CH). Before moving to Berlin, she worked as an associate researcher in the Developmental Robotics Group at Sony Computer Science Labs in Paris. She has published more than 50 papers in renowned scientific journals, books and peer-reviewed conference proceedings. She is PI in two DFG graduate schools and PI in the EU Initial Training Network INTRO on interactive robotics. Her research interests include sensorimotor interaction and learning, joint attention, and intuitive human-robot interaction.