ELSEVIER

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot



Telenoid android robot as an embodied perceptual social regulation medium engaging natural human-humanoid interaction



Rosario Sorbello^{a,*}, Antonio Chella^a, Carmelo Calí^b, Marcello Giardina^a, Shuichi Nishio^c, Hiroshi Ishiguro^{d,c}

^a Dept. DICGIM, RoboticsLab, Universitá di Palermo, V. delle Scienze, Palermo, Italy

^b Dept. Scienze Umanistiche, Universitá di Palermo, V. delle Scienze, Palermo, Italy

^c Hiroshi Ishiguro Laboratory, ATR, 2-2 Hikaridai, Keihanna Science City, Kyoto, Japan

^d Department of Systems Innovation, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka, Japan

HIGHLIGHTS

- We present an analysis about the main features of human-humanoid interaction.
- We conduct an extensive test with 142 people employing the Telenoid android robot.
- The Telenoid is perceived as a cooperative agent for a shared environment.

• Perception and believability make the Telenoid a socially acceptable robot.

ARTICLE INFO

Article history: Available online 30 April 2014

Keywords: Telenoid Geminoid Social robot Human-humanoid robot interaction

ABSTRACT

The present paper aims to validate our research on human-humanoid interaction (HHI) using the minimalist humanoid robot Telenoid. We conducted the human-robot interaction test with 142 young people who had no prior interaction experience with this robot. The main goal is the analysis of the two social dimensions ("*Perception*" and "*Believability*") useful for increasing the natural behaviour between users and Telenoid. We administered our custom questionnaire to human subjects in association with a well defined experimental setting ("ordinary and goal-guided task"). A thorough analysis of the questionnaires has been calculated. Our experimental results show that the perceptual behaviour and believability, as implicit social competences, could improve the meaningfulness and the natural-like sense of human-humanoid interaction in everyday life task-driven activities. Telenoid is perceived as an autonomous cooperative agent for a shared environment by human beings.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Since humanoid robots are going to be part of the life of human beings, specific studies are oriented to investigating collaborative and social features related to human-humanoid interaction (HHI) [1–3]. The HHI is oriented nowadays towards a cohabitation environment where human and humanoid will share common tasks and goals [4]. In particular Kanda et al. [5] focused their attention to the concept of "communication" humanoid robot thinking as a partner to facilitate some human activities. Oztop

et al. [6] put their attention to understand the perceptual relation between human and humanoid robots. The iCat developed by Poel et al. [7], is a user-interface robot able to display a range of emotions through its facial features and it is mostly controlled by predefined animations. The ICub [8], is a child humanoid robot that it is used in embodied cognition research. In contrast to these typical humanoid robots, Geminoid HI-1 is a humanoid robot with the external appearance of its inventor, Prof. Hiroshi Ishiguro and it is thought to be indistinguishable from real humans at first sight [9–11]. In particular, much relevant literature appeared on the features of the natural role of agent interaction [12,13]. The minimal agency includes a key aspect that is defined as "sense of co-presence" [14], [15]. We oriented our research in the HHI field in the direction of "the sense of being together with other people in a shared virtual environment" [16]. In particular we are interested

^{*} Corresponding author. Tel.: +39 3289859060.

E-mail addresses: rosario.sorbello@unipa.it, sorbello.rosario@gmail.com (R. Sorbello).



Fig. 2. The two stages of interaction setup.

in the study related to the sense of a person to be present in a remote environment with a robot ("Telepresence") and to the sense of a person to be present in a common environment with a robot where humans and humanoid are "accessible, available and subject to one another" [17]. In this paper we introduce our research that aims at investigating the social-cognitive and the underlying perceptual skills, which are likely to contribute to the sociability in the human-humanoid interaction (HHI). We construed perceptual and social abilities as indicators that allow to assess the nature of interaction from the human agent's point of view, the extent of recognition of the humanoid robot as an actual agent in free and task-driven contexts of interaction. The assumption is that the HHI is a context of interaction in which the perceptual and social-cognitive abilities of ordinary day-life are specialized. Therefore, the naturalness and the efficacy of the HHI interaction depend on how much human agents can recognize, albeit implicitly, and exploit these skills which correspond to our indicators. We arranged the indicators into two distinct constructs, according to their mainly perceptual or social nature, whose subdimensions correspond to the perceptual and social-cognitive skills which make ordinary interaction effective in daily life. Then, the constructs were modelled on the structure of a questionnaire, which was administered to subjects in controlled interaction conditions. Hence the indicators were formulated as various items of a structured questionnaire. As a consequence, the degree of naturalness and efficacy of the HHI interaction is considered equivalent to the favourableness that is measured on the ground of the sum of subjects' scores of agreement on an ordinal Likert scale. This paper does not only deal with the principles and assumptions underlying the constructs, rather it specifies the perceptual and social-cognitive sub-dimensions of the two constructs, it describes how the indicators work if embedded in an HHI set with the Telenoid robot, finally it provides a complete descriptive analysis of the results.

2. Telenoid robot

Telenoid, as shown in Fig. 1, is a teleoperated android robot [2] with a minimal human likeliness design that can resemble anybody. Telenoid was created by choosing features useful for

communication with humans and eliminating the non-neutral ones. Due to its minimal design, it allows people to feel as if a distant acquaintance is actually close.

A GUI button or a GUI tablet controls the specific movements of the arms and head to remotely embody the operator's behaviours and emotions. The aim is to create a minimal human embodiment that allows any individual to transfer her/his own presence or, better, to visualize it in a distant location by mediation. Our research is primarily centred on the human agent's point of view in order to investigate whether the humanoid robot is considered not merely as a monitor or a screen on which another agent is projected rather as an artificial but at the same time actual agent whose behaviour displays the sense of co-presence which contributes substantially to the experience of a meaningful and effective interaction. Our research deals with the perceptual and social-cognitive abilities underlying the HHI. Therefore Telenoid's movements provide a suitable test bed for a minimal set of the perceptual's and social-cognitive's abilities.

3. The proposed approach

In the *Experimental Setup section* we describe the phases of the interaction scenario, the characteristics of the participants' samples, the perception and social-cognitive constructs of the questionnaire. In the *Questionnaire section* we describe in detail the constructs and their sub-dimensions.

3.1. Experimental setup

Students of the Faculty of Architecture and Engineering (University of Palermo) who did not have prior interaction experience with humanoid robots were recruited for the tests. All participants (142 total, 85 males and 57 females) have been introduced to the Telenoid, to the interaction setting structure with the robot that required a two stage interaction, as shown in Fig. 2, and to fill in a questionnaire. All interactions were videotaped.

A first free interaction stage was meant to allow subjects to adapt either to interact with the humanoid robot or to acquire, as early as possible, the skills for operating the robot through the

Table I	
Subject sample	es characteristics.

T 11 4

Subjects:		
-Male	59.85%	
-Female	40.15%	
-Average age	21.77	
Previous knowledge or general acquaintance and robotics issues		
Personal attitude to robotics		
-Real interest	51%	
-Significant knowledge	9%	
-Curiosity	37%	
Degree of agreement on the Acceptance of Robots in the near Future		
-Accept as useful tools in everyday life	73%	
-Accept as useful tools in jobs	69%	
-Accept as companions in the other's life	47%	
-Accept as companions in one's own lifetime	51%	
-Accept with suspicion	28%	

Table 2

The Perception construct and the questions modelled after its sub-dimensions.

Perception construct (P)

Sub-dimension P1: Sense of HHI and of shared environment

- P1.1 At which distance does a successful interaction obtain?
- P1.2 Which activity did you perform to obtain a face value optimal interaction after taking the robot away from its initial position?
- P1.3 Which is the description that fits how you felt the interaction set up?
- P1.4 What should you do to make the interaction more effective and understandable?
- Sub-dimension P2: Perceptual clues of apparent behaviour
 - P2.1 Which part of the Telenoid's face caught your attention?
 - P2.2 Is the coupling between the Telenoid's head movement and voice reproduction natural-like and effective?
 - P2.3 Is the Telenoid's gaze focused on you during the interaction?
 - P2.4 What does the Telenoid's gaze make it look like?
 - P2.5 How much is the Telenoid's eye-head-lips coordination consistent with discourse parsing and turn taking?

Table 3

The believability construct and the questions modelled after its sub-dimensions.

```
Believability construct (B)

Subdimension B1: Valence

• B1.1: What makes the Telenoid interact?

Subdimension B2: Motivation

• B2.1: What is the apparent motivation of the telenoid's interest in the

interaction?

Subdimension B3: Value

• B3.1: Why does the Telenoid tune its behaviour to yours?

Subdimension B4: Naive reason of reliability

• B4.1: The Telenoid looks reliable when?

Subdimension B5: Social Attitude

• B5.1: How does the Telenoid's behaviour look like?
```

control box. A second interaction stage was instead task driven. Participants were allowed to choose one interaction scenario among those which were available that ranged from the ones related to: booking a hotel reservation, making a phone call to a mobile company to obtain a contractor services information, to matriculate or to enter his/her name or one of his/her fellow ones for a course examination by talking directly with the robot. Table 1 illustrates the subject sample characteristics, the degree of familiarity with robots and of robots acceptance.

Table 2 and Table 3, show the constructs (*Perception* (**P**) and *Believability* (**B**)), their sub-dimensions and the questions used for the layout of the questionnaire.

3.2. Questionnaire: constructs and sub-dimensions

The construct Perception was built according to a cognitive interpretation of perceptual abilities, that is to say the perceptual aspect of the interaction was considered not primarily to verify the embodiment of the robot, rather to investigate which perceptual features are used by human agents to understand the overt behaviour of the robot in the course of the interaction. Therefore, we decomposed the construct Perception into the following two subdimensions: sub-dimension P1 which refers to the sense of HHI on the ground of phenomenal distance and shared environment and sub-dimension P2 which refers to the apparent behaviour. It is composed by a set of perceptual clues which on the one hand allow agents to understand behaviour, by coupling effective actions with meanings and intentions, and on the other hand enable a humanoid agent to appear as an actual agent rather than solely as an artificial intermediary between humans. We reasoned that the issues of HHI can profit from the debate on mind reading and behaviour understanding abilities, in particular in connection with the discussion on the theory of mind, which is based both on theoretical arguments and experimental evidence [18,19]. We derived the principled assumption that the interaction among agents, biological or artificial, always occurs in an interactive space whose distances and regions are fixed in terms of possible actions and effects from the agents' point of view. A noteworthy implication is that it is reasonable to claim that agents recognize mutually as such and, in the particular case of humanoids, robots' behaviour is taken as meaningful and effective if some perceptual clues are coherent and consistent. The construct Perception was designed accordingly. The P1 sub-dimension concerns first the distance as a perceptual feature of the interaction and the environment where the interaction obtains. The distance is meant not as metric rather as proxemic and, in particular, as it looks like from the human agent's point of view. The value it holds as a perceptual clue for agents depends on the possibility it grants them to focus on perceptual clues, which may convey the intentions and meanings. Second, P1 concerns how the environment is perceived, which is a feature that can determine the role agents play in it. The P2 sub-dimension regards clues for understanding behaviour, which are provided by the perception of face, gaze focusing and gaze contact, as well as the apparent mutual coherence of eye movements with head movements and voice sound emission. These clues could affect the attribution of intentions, the discourse parsing and how agent's general attitudes to interaction is inferred. In conclusion, P1 concerns the perceptual space of the interaction, P2 concerns the relevance and coherence of a minimal set of perceptual clues by means of which a humanoid robot can be considered as a real agent. The construct Believability was designed on the basis of the relevant literature to the social character of HRI [20]. The discussion in literature focuses on highlighting the social-cognitive skills which allow robots to obtain a useful and effective social interaction with man. In accord with the approach of the first construct, we tried to specify the sub-dimensions of the construct Believability in such a way to link the social abilities to the extent at which they can be displayed in the robot's apparent behaviour. The sub-dimension B1 concerns whether human agents perceive the Telenoid to have a sort of value system on whose ground it may consider suitable to interact with a human agent. The sub-dimension B2 concerns whether human agents recognize the Telenoid as able to display some social capability which motivates its behaviour. The sub-dimension B3 is about whether human agents take the Telenoid as displaying a value for tuning its behaviour to them, hence which value may it be. The sub-dimension B4 concerns the reliability of the Telenoid. In other words, B4 is about whether human agents use a socialcognitive yardstick to judge the reason why the Telenoid seems to tune its behaviour to theirs and if this reason may be considered as a sort of determinant of the Telenoid's behaviour in the interaction. In agreement with this approach, the research is mainly centred on the human agent's point of view in order to investigate whether the humanoid robot is considered not merely as a monitor or a screen on which another agent is projected rather as an artificial but at the same time actual agent whose behaviour displays the sense of co-presence which contributes substantially to the experience of a meaningful and effective interaction. The perceptual indicators that are certainly the substantial element of our research were crucial with no doubt starting from the grades of freedom of movement (DOF) of the Telenoid as well as their demonstration in the course of interaction depending on the decisions and actions carried out by the human in charge of tele-operating the robot. The tele-operating aspect of working on the robot was not underrated nor missed. We provided a further section of our guestionnaire solely to subjects who had tele-operated the Telenoid. However, our interests from a technical point of view are addressed primarily to the part that concerns the necessary and effective control conditions by the robot's tele-operator in order to let the Telenoid perform its behaviour which should make the interaction with human agents successful. In the same way, the Telenoid's movements of the eyes, mouth, head and arms provided a suitable test bed for a minimal set of the perceptual and socialcognitive abilities, which are essential for any effective instance of interaction to occur, either in human-human everyday-life or in human-humanoid specialized condition. On the basis of the tele-operator's evaluation of the Telenoid's DOF suitability and its movements to display the perceptual indicators, which are strictly connected to an effective social behaviour, we wanted to know how much the Telenoid could appear independent in the human agent's eyes. In the future research we aim at introducing some degrees of semi-autonomy, which will make the robot even more reliable in the course of interaction, providing it with a control on the attention of the interacting human agent, with topic recognition and question answering skills. Furthermore, a promising research area can be that of improving the capacity of the robot to display explicitly the emotions of the tele-operating subject to create a natural-like context of interaction in which the use of the social-cognitive abilities can be more easily specialized.

4. Experimental results

We obtained results by having subjects answer the questionnaire which was structured according to the sub-dimensions of the perceptual and social-cognitive abilities of the constructs. The questionnaire was composed of single forced choice and five points Likert questions. We treated Likert item scores as ordinal data. Accordingly, we visualized subjects' responses by means of bar charts and, for the Likert questions, by means of a box plot representation which allows to evaluate central and dispersion measures of ordinal data.

4.1. Evaluation of the reliability and the internal consistency of the questionnaires

The constructs were formulated as the items of the questionnaire which was administered to the subjects. The questionnaire is composed of single-forced choice and five point Likert questions. Questions represented the distinct sub-dimensions of the constructs.

Once the subjects' responses were codified, a standard analysis of internal consistency and reliability was performed on these data.

The questionnaire is assumed to provide information about the two constructs if in the sum of the subject's response scores the error is minimized in contrast to the values that measure the latent sub-dimensions of the constructs. Calculating Cronbach's alpha is a standard method to assess the internal structure of the questionnaire with respect to the relationship between the variance for each response score and the variance of their sum [21]. This relation gives us an estimation of the proportion of the true score, **viz**. the values of the responses which measure the constructs. The greater the variance of the sum is as compared with the sum of the variance of each response, the more the questionnaire is internally consistent.

Since we designed a questionnaire for two constructs, we calculated Cronbach's alpha for both sections of questions for each of the two constructs. The alpha value is ≈ 0.88 for the items of the questions grouped under the construct Perception. The alpha value is ≈ 0.77 for the items of the questions grouped under the construct Believability (see [22] for the acceptable thresholds for alpha).

As an additional indicator of reliability, we computed how much the Cronbach alpha values were a consistent assessment of the two constructs of interest by means of the Pearson coefficient of correlation. Given the nature of the responses in our Likert scale questionnaire, we used the odd–even split-half method in order to reduce the component of uncertain variance in the two sums, for each of the two constructs. The reliability of the questionnaire is assumed to correspond to the inter-correlation of response scores for the items of the questions in either section of the questionnaire. For the items of the construct Perception the Pearson coefficient value is ≈ 0.97 . For the items of the construct Believability the Pearson coefficient value is ≈ 0.59 . Finally, we corrected these measures by the Spearman–Brown formula and obtained the value of ≈ 0.98 for Perception and the value of ≈ 0.74 for Believability.

4.2. Analysis of questions responses

The first two questions (Figs. 3 and 4) concern the distance as a perceptual feature that contributes to establishing the sense of interaction.

P1.1: Which is the best distance to keep in order to obtain a successful interaction?

Very few subjects (4%), as we have shown in Fig. 3 report to prefer the Telenoid to remain in its initial standard position as well as a small number of them claim the Telenoid's position, hence its distance and allowance into a subject's interaction space, What distance should a man keep in order to obtain a successful interaction?





Which action did you perform to obtain a face value optimal interaction after taking the robot away from its initial position? 60.00% 50.00% 50.00% 38.10% 40.00% 30.00% 20.00% 11.90% 10.00% 0.00% No preference about distance with telenoid I preferred to get closer to telenoid with I preferred to push away the telenoid with respect to its closer initial respect to initial position



Which is the description that fits how you felt the interaction set up?





to be indifferent. Higher percentages are reported for selecting a distance that allows the robot to access the inner zones of a subject's personal space. The preference is given to the near distance of the face position (30%), which presumably is connected with the possibility of looking for relevant information directly in the robot's face where most ordinary clues for intention reading are usually located, as well as to the "at arm's length" position, which presumably allows subject to have as it were at a glance such perceptual clues as those provided by the eyes in connection with head and arm movement. It is noteworthy that given the weight of the robot a more physically comfortable position would have been located at subjects' knees. Actually 18% of our subject reported to prefer this distance.

P1.2: Which action did you perform to obtain a face value optimal interaction after taking the robot away from its initial position?

Half of our subjects (50%), as we have shown in Fig. 4, report to have preferred letting somehow the Telenoid getting closer to them after taking the robot from its initial standard position, in order to have an optimal interaction with the robot. Only 11% report to have preferred pushing it away. This finding can be connected to the responses to the previous question. If the distance is a perceptual feature of the interaction space, rather than a metrical property, its contribution to the sense of interaction comes about by adequately and pro-actively selecting it. It is interesting however to highlight the response of 38% of the subjects who claim to have no preference in selecting somehow a distance to optimize the interaction with the Telenoid. It could mean indifference to the alternative which is given in the question rather than to the issue itself.

P1.3: Which is the description that fits how you felt the interaction set up?

This question concerns the perception of the interaction space where HHI occurs. To assess the alternative views on the interaction space, which can be either accessed from distinct reference points or felt as a sort of shared environment, these views were formulated as distinct items about which subjects were asked to score their rate of agreement. 68% of subjects, as we have shown in Fig. 5, agree or strongly agree that they felt the interaction set up as a shared environment where they and the Telenoid act on a par. Only 22% disagree and the percentage of strongly disagreement is very low (9%). 37% of subjects agree that they felt to be the reference point of the environment in which the interaction occurred. This seems plausible in the light of the task-driven part of the interaction set up, which was designed so that the Telenoid was required to meet subjects' requests. It is then noteworthy that in comparison with the first item the score of strongly agreement decreases, while the rate of indifference raises and almost doubles (from 22% to 40%), which can be a sign of a high likelihood that this description is not meant as adequately fitting the perception of the interaction space. This seems to be the case for the third item which proposed the description according to which it was the Telenoid to be felt as the reference point of the interaction space.



Fig. 6. Question P1.4.





Fig. 7. Question P2.1.

42% rate of indifference means that regarding the Telenoid as the access point to the environment of interaction it is actually not at all discriminating rather than misleading or openly wrong. This interpretation is justified by the fact that almost the same number of subjects (a fourth of them) either agree or disagree.

P1.4: What should you do to make the interaction more effective and understandable?

The items of this question are formulated to investigate how subjects approach the perceptual options, which are available to them to make the interaction effective, according to the sense of the interaction space. 50% and 55%, as we shown in Fig. 6, of subjects agree, correspondingly 11% and 10% strongly agree, that looking at and touching the Telenoid at a variable distance helps understand the robot's behaviour. This estimate is meaningful because it refers to the fact that among the allowed sensory modalities the most important one for the human user is the selection of the optimal distance in order to allow a long lasting and successful interaction. This approach can be consistent with the sense of a shared environment in which the agents act on a par. 34% agree and 8% strongly agree that a feasible option amounts to looking at the robot from a fixed distance. 23% agree and 6% strongly agree that touching the robot at a fixed position is a feasible option. The difference in agreement can be due to the distinct sensory modality: visual perception accommodates better with a fixed standpoint, albeit momentarily, than tactile perception. However, these two items which refer to the persistence of the standpoint report either an increasing rate of disagreement (29% and 32% respectively) or an increasing rate of indifference in the case of touch, which at 37% is the highest score for the item. It is to be highlighted however that the same rate of indifference (25%) is obtained by looking from both a fixed and a variable standpoint for the visual modality, even though this score is given within two different overall response score profiles. The next group of questions is about the perceptual clues of apparent behaviour that are ordinarily used to understand the other agents' intentions.

P2.1: Which part of the telenoid's face caught your attention?

It is well known that face perception is essential in understanding behaviour. This question addressed the issue of which part was considered to be more informative, albeit implicitly, by human agents in the course of the interaction, since that the upper and the lower part of the face seem to convey distinct information. 50% of subjects, as we have shown in Fig. 7, paid attention to the eyes of the robot, while 37% to the whole face. It is not surprising that subjects attend to such a meaningful clue both in everyday life and in this setup. The attention to the whole face attests the will to get much of the information, which is available in the other agents' face as it were at the same time. It is striking that only 6% claimed to have attended to the mouth. One of the Telenoid's effectual feature is the lip movement which is synchronized with the sound reproduction of the tele-operator's voice. It is likely that for technical reasons due to the elasticity of the silicon skin on the robot the lip movements are not so manifest to agents who interact with the robot for the first time.

P2.2: Is the coupling between telenoid's head motion and voice reproduction natural-like and effective?

The range of the distribution shows that subject scores are nearer to one another than the responses to other questions on similar issues about the other perceptual clues. The central 50%, as we have shown in Fig. 8, is concentrated on the positive side of the agreement rate. Actually, 44% agree and 25% strongly agree (not shown) that this coupling is perceived as natural-like and effective. The median location and the relative index of dispersion show that the distribution is negatively skewed.

P2.3: Is the telenoid's gaze focused on you during the interaction? This question tries to investigate whether subjects perceived the Telenoid's gaze as able to be focused that is a prerequisite of such a meaningful perceptual clue for understanding the role, which agents can play in any interaction, as the gaze contact. If compared to the scores of the previous question, the increased difference between the first and the third quartile (Inter-quartile Range) shows that the dispersion of the 50% of central rating scores is increased with reference to the median. The figure of the distribution is symmetrical. Actually, 18% of subjects, as we have shown in Fig. 9, neither agree nor disagree that the gaze focus is perceivable despite that 41% of subjects agree and 15% strongly agree (not shown).

P2.4: How is the telenoid's gaze affected in relation to its behaviour?

This multi-item question, as we have shown in Fig. 10, is meant to cover which is the effect, if any, on how the gaze of the robot





Fig. 11. Question P2.5.

makes it look like with respect to its general attitude towards the interaction with human agents. The positive item of looking like expert covers a spread range. The Inter-quartile range shows that the 50% of the score are less dispersed with reference to the median score in the positive ratings of agreement, besides the scores of the last upper quartile are more bunched than those of the first lower quartile (14% strongly agree and 9% strongly disagree). The distribution is positively skewed. Subjects were asked also to claim their agreement on negative items. The acquiescence and hostility of the Telenoid as the attitude, which could be carried by its gaze, score a low rating of agreement. The central 50% of responses is located below the indifference middle point, the inter-quartile range show that they are concentrated around the low median with negative scores bunched in the first quartile. As to the apparent indifferent attitude of the robot, the dispersion of the 50% ratings and the fact that the coefficient of quartile deviation is greater than for the former two negative items could

6 5 4

attest that this item was not feel as much as discriminating by the subjects.

of the user

P2.5: How much is the telenoid's eye-head-lips coordination consistent with discourse parsing and turn taking?

This multi-item question, as we have shown in Fig. 11, is meant to investigate at which extent, if any, the perceived coordination of the eye, head and lip motions of the Telenoid can be used as a perceptual clue for discourse parsing and turn taking, which are abilities that have been emphasized as substantive aspects of a successful interaction also in virtual environments or verbal interaction with artificial agents. This coordination seems to be not very effective in acting as such a perceptual clue for these demanding cognitive skills. 50% of the scores with reference to its use as a signal of the beginning of the discourse is dispersed and positively skewed. In the other item ratings, the distribution tends to become more symmetrical and has virtually the same coefficient



Fig. 13. Question B2.1.

of quartile deviation. The following questions are about the subdimensions of the Believability construct.

B1.1: What makes the telenoid interact?

This multi-items question, as we have shown in Fig. 12, is about the valence of the Telenoid as it can appear in the course of the interaction. Intuitively, subjects are asked about whether they perceive the Telenoid's source of action. Does it depend on a sort of value system, which somehow suggests it to interact? If any, is it internal or external?

Low scores of agreements are shown by the item that the Telenoid's behaviour appears to depend on a sort of inner value system according to which it could be able to tell what he likes from what he dislikes. The central 50% is concentrated on the lower side below the indifference middle point. Most scores have values that are lesser than the median and greater than those of the respondent scores of the first quartile. A similar picture is given for the item that the Telenoid reacts to the human agent's behaviour on the ground of its evaluation of what happens in the external environment. This is not such a surprising result, given the design of the interaction and what subjects actually knew about the Telenoid's real working.

The item fares better in subjects' rate of agreement that the Telenoid reacts according to the human agent's behaviour. The distribution is still positively skewed and the range is more dispersed, however 50% of the responses shift to the positive side of the scores, data are more bunched in the upper last quartile than in the lower first.

The last item that the Telenoid looks like to behave in order to give a correct answer to the human agent's requests seems to be felt as not discriminating, since the distribution and form of the scores show a flattening of the frequency of responses across the various rating modalities, with a greater coefficient of quartile deviation. The characterization of the answer as the correct one, which the robot is deemed to give, is taken by the subject as requiring a more demanding and complex system of decision making and evaluation that they know the robot cannot have. B2.1: What is the apparent motivation of the telenoid's interest in the interaction?

This multi-items question, as we have shown in Fig. 13, is about the motivation of the Telenoid's behaviour during the interaction. Intuitively, subjects are asked about what moves the robot in displays interest in the requests of human agents. Subjects' ratings score agreement on the apparent ability of the Telenoid to pay attention, show to be interested in and looking like competent, notwithstanding the noticeable difference that the responses to the latter two items show a greater range dispersion, even though the inter-quartile range is the same. The distribution of the central 50% of responses for these three items is negatively skewed. The responses to the item that the Telenoid shows to be motivated by its ability to understand the human agent's needs show a greater frequency of scores which are lesser than the median and greater than those of the first quartile. This result is explained by the fact that there is an increase of the percentage of the disagreement rating as well as of the indifference rating, though in this case it is more moderate.

B3.1: Why does the telenoid tune its behaviour to yours?

This multi-items question, as we have shown in Fig. 14, is about the value of the Telenoid's behaviour under the particular respect that it appears to tune its behaviour to human agents. Intuitively, subjects are asked about whether they recognize an apparent value for Telenoid to coordinate its behaviour to theirs? If any, which is this value? The items that regard the Telenoid to attach value to cooperate, being friendly and fair with the human agents record a positive agreement rating. With reference to the central 50% of score responses, the items record much more frequent responses which have a higher score than the median. For the item that qualifies the Telenoid as somewhat appreciating being fair, the agreement reaches such percentages of 65% of agree and 13% of strongly agree (percentages not shown). As to the item that the Telenoid considers to be important to give the most suitable answer to human agents, the range increases though the distribution of the central 50% does not change, the lower first quartile data become much more spread and, above all, the





Fig. 16. Question B5.1.

distribution inverts the skewness, which is due to a significant increase of indifference ratings.

B4.1: The telenoid looks like reliable if:

understand

This multi-items question, as we have shown in Fig. 15, is about the reliability of the Telenoid as it can be assessed in so far as subjects can identify a yardstick to judge it on the ground of its overt behaviour. Intuitively, subjects are asked on whose apparent grounds they feel comfortable in attributing reliability to the robot. Subjects incline to not give their full agreement to the fact that the Telenoid can look reliable if it shows to feel autonomously emotions. This does not appear to be a sensible determinant of its behaviour and cannot count as a yardstick of reliability. The central 50% is shifted to the low ratings of agreement, the distribution of the response scores is almost symmetrical, the lowest quartile data are bunched. Nor do they agree that the Telenoid looks reliable because it shows the ability to match the information about its one's own and human agents' states. The items fare better scores which propose more overt characteristics as yardsticks of reliability such as being able to react as to satisfy the human agent requests or to adapt its own behaviour according to the interaction needs. The inter-quartile range decreases, the central 50% is shifted towards more positive agreement ratings, the skewness of the distribution is reversed.

B5.1: What does the telenoid's behaviour look like?

Finally, this last question, as we have shown in Fig. 16, asks subjects to rate their overall agreement on the behaviour of the Telenoid in connection with their judgements on some social attitude, which ordinarily can contribute to improve the way agents interact with one another. Subjects' responses emphasize the enjoyable character of their experience as it can be easily determined by the distribution of the scores to the item according to which the behaviour of the Telenoid is funny. It is noteworthy that this description does not hinder subjects to acknowledge some social attitudes, which contribute in making the interaction efficient and successful. Honesty is rated with high scores of the agreement. The central 50% is concentrated above the middle point and the distribution is negatively skewed. The ease of understanding the robot's behaviour has the same pattern of response scores. The attitudes of sociability and efficacy have a larger range variation, but the inter-quartile difference and the coefficient of variation are the same as the former two items. The same holds for the inspiring trust character of the Telenoid's behaviour, although the agreement scores are positively skewed. The item about the social ability of appearing persuasive does not score rates of agreement. 75% of the responses is below the middle point of indifference. A plausible explanation of this score



Fig. 17. Question: D1, D2, D3.







may be a certain amount of ambiguity in the meaning of the word "persuasive". A further clarification may be required to factor the component meanings of being plausible, cogent, convincing, impressive.

4.3. Analysis of questions on the teleoperation side of interaction

The following set of questions, as we have shown in Table 4, was designed to gather information on the judgement on the usability of the Telenoid with particular reference to the connection between the technical characteristics of the control and the possibility of letting the Telenoid behave in such a way to match the tele-operator's presence in the robot's overt behaviour and the expectations of the users on what make the behaviour understandable and effective.

The first multi-item question, as is shown in Fig. 17, concerns the capability and the ease to match some intended perceptual clues with the actual motions and aspects of the Telenoid in order to let the human agents gain some useful information on the ongoing interaction. The items score a sufficiently positive agreement, which are about the ability to reproduce head movements, which are coupled to the perceptual clues that are retrieved in the robot's face, and in general to visualize one's own intention as to the successful development of the interaction. 50% of the scores, as we have shown in Fig. 17, is above the middle point of indifference, its distribution is not spread, and most frequent responses are given which are higher than the median with overall 75% of responses which lies on the positive side of the agreement. The item fares worse which is about the ability of the Telenoid's motions to visualize the tele-operator's emotions. Scores are shifted below the indifference point.

D4: Did you notice a delay among the voice and motion commands and the telenoid's reaction?

This question D4, as we have shown in Fig. 18, means to acquire some information about what could disturb the matching between the tele-operator's intentions and its visualization for the human users during the interaction.

D5: Does the overall view of the interaction space fit your needs for a successful interaction?

This question D5, as we have shown in Fig. 19, is meant to gather information concerning the view that the control devices allow the whole interaction space. This question is then ideally coupled with one of the aspect of the sense of interaction from the human agent's point of view, that is to say the perception of a shared environment of the interaction. Subjects highly agree that the direct view of the user is effective for the matching and visualization of intentions to obtain, which is likely required for a successful interaction, and above all for the ability of adjusting the Telenoid's behaviour as the





Fig. 19. Question: D5.

interaction goes along. The issue about the efficacy of the direct view of the overall environment is not as good. The dispersion is around the median, which is measured by the coefficient of quartile deviation is greater, though 75% of the scores are located in the regions of agreement or strongly agreed.

6

5. Discussion

Our research deals with the perceptual and social-cognitive abilities underlying the HHI. As far as the social aspect of interaction is concerned, our assumption agrees with Dautenhahn [23] who emphasizes that social intelligence is an essential aspect of human intelligence. As a consequence, in order to build human-like and believable robots it is required to take the social intelligence into account as a fundamental prerequisite for any artificially intelligent robot. However, we try to specify further these social-cognitive abilities in connection with the perceptual abilities, which are the perceptual basis on whose ground the robot's behaviour can become understandable and believable. Therefore, the issue of the sociability is restricted to how agents perceive the apparent behaviour of the humanoid robot. The present research abides by the fundamental assumption of the human-centred HRI, according to which the robot has to be believable, acceptable, comfortable to humans as well as appearing to share a common environment with them. Since we are interested in the social acceptability of the humanoid robot in the context of an efficient interaction, the perceptual basis of behaviour is not restricted to the exterior aspect of the robot [24]. The Perception construct is meant to investigate the cognitive means by which agents are able to understand behaviour as it is displayed during the interaction. The social-cognitive dimension of the interaction was dealt with in the work of [25] who emphasizes that the social-emotional intelligence is a useful means for understanding the behaviour. However, this work assumes that mental social models are required to attribute mental states to other agents, while we suggest that any agent's behaviour has to satisfy in a consistent and coherent way the perceptual features and clues, which are already effective in everyday-life ordinary contexts. This argument touches upon the heavy debated issues of the theory of mind as an explanation of the folk psychological ability of other minds reading [19]. Given that the mental states are not inferential constructs that are derived by means of a naïve inner theory of mind. Gallese et al. construe the experimental evidence on the so-called neurobiological mirror system such that the conclusion is drawn that mental states are attributed to other agents if their motor behaviour corresponds to what the subject wants to do and experiences if he/she had the same intention. Therefore, the mind reading would amount to simulating other agents' apparent behaviour which becomes meaningful once it is as it were transposed within one's own

point of view. Graziano [26] shows how socially meaningful information is encoded in neurons which respond selectively to movements, actions of agents in a spatial multiple coordinate system that is centred in agents' body parts. From the overt behaviour's standpoint, the conclusion can be made that the spatial frames at arm's, limb's, hand's, face's length are as many as maps for intentions-actions meaning. Gallese [18] argues that recent neuroscientific findings point out that understanding behaviour and attributing mental states is founded on a multi-level experience of how agents' bodies interact in a shared space. Some specified neuronal systems allow social ontology and action to be mapped in "social perception" by which intentions of other agents are conveyed in the phenomenal content of perceptual behaviour. Social cognitive mental skills allow subjects to retrieve mental contents of others by means of the features of apparent behaviour. Calí [27] suggests that these arguments and evidence can be construed as implying that a direct (and not knowledge-dependent) access for agents obtains to their shared phenomenal content of behaviour. This access is essential to understand the intentions and actions of other agents on the ground of the features of apparent overt behaviour. Some of these features can be exploited by agents as they were a perceptual mapping of behaviour on whose ground a functional equivalence between the mental states of the observer and of other agents' is extracted through the interaction. The main difference between agents would primarily rest on the different locations of their points of view from which a shared environment is accessed. On this basis, our assumption is justified according to which the perceptual appearance is not merely the outward aspect of the robot [24]. Rather it plays a substantive cognitive role, which distinguishes our approach from those characterized by the opposition appearance vs. reality [25].

6. Conclusion and future works

In this paper the research aims at finding whether perceptual and social-cognitive abilities, which underlie ordinary day-life, can be specialized and exploited by agents in HHI contexts to make the interaction natural-like and effective. The assumption was made that the perceptual indicators and the social cognitive abilities constitute a sort of implicit specialized competence which contributes to a successful interaction with humanoid robots. Then, the specification of the perceptual and the social cognitive indicators for interaction can provide a principled approach for the study of the cognitive characteristics of HHI and pave the way for exploiting the potentiality of these clues to improve the naturalness and efficacy of interaction with robots in the future. This research abides by the view that the sense of "togetherness" between persons and Humanoid considered as Robot Agent is "inherently social" and is highly connected with the concepts of particular behaviour defined "sensible" because [4] capable to express cognitive functionality. We built two constructs able to cover these perceptions and social-cognitive indicators in order to transpose some of these cognitive dimensions of ordinary everyday life and to embed them in a controlled interaction set up. By administering a questionnaire we wanted to discover whether and at which extent some clues of these dimensions contribute to the originating of the intentions and meanings of a humanoid's behaviour provided which should be prepared to make it look like consistent and coherent from the human agent's point of view with whom interaction occurs. The humanoid robot Telenoid was used because its technical and apparent characteristics make it suitable to carry out our research. According to the present stage of our research, this study is to be sure still exploratory. Nevertheless, some findings emerge which emphasize the role played by the perceptual and social-cognitive clues in making the interaction acceptable and efficient in task driven contexts. We found that the sensation of sharing a common environment is of essential importance to the view that interaction occurs in a space in which human and humanoid agents are on a par. In this interaction space each participant selected his/her own distance in order to exploit usual perceptual clues to read off the meaning of interaction from apparent behaviour. We found that such clues provided by face, gaze, eye-head-voice coordination are perceived as meaningful by human agents, even though they are not sufficient to transpose a demanding task such as for instance discourse parsing. The present research does not restrict the study of the perceptual basis of interaction to the outward appearance of the robot. Accordingly, the sub-dimensions of the Believability have been investigated and so far chosen as they are displayed in an apparent behaviour. Notwithstanding a full interpretation of all the results for the Believability items may require further clarification. An important result is that subjects significantly inclined to perceive the Telenoid as cooperative and competent event though this favourableness seems to be not associated with the use of a clear cut vardstick to assess the reliability of the robot. Further research is needed to take into account a different characterization of the various meaning such a vardstick may assume, its robustness across various tasks and distinct interaction contexts as well as in experimental conditions where subjects do not actually know how the Telenoid really works. The design of conditions which allow a longitudinal study of subjects' responses may be further needed. Though exploratory this research stage may be, evidence can be found that subjects who are involved in experiments with the robot do not seem to need such a strong yardstick, at least insofar as it is intended as an external standard to assess the reliability of the robot that would compel the attribution to the humanoid robot of an inner determinant for its behaviour. Provided that perception and social-cognitive clues are consistent and fit the interaction task, subjects favourably inclined to accept the humanoid's behaviour as tuned to their requests, albeit within the bounds of our controlled interaction, and motivated by the commitment to satisfy their requests. It is sufficient that the Telenoid shows to be able at a certain degree to satisfy the human agent requests and to adapt its own behaviour according to the interaction needs. The Telenoid also appears to embody such valuable traits as being interested, friendly, fair and in taking care of subjects, which help in making the task driven interaction successful. It is to be noted that more demanding clues seem to be not supported by current Telenoid's apparent behaviour. The questionnaire, which was modelled after the constructs sub-dimensions, is reliable and internally consistent. The complete descriptive ordinal analysis of subjects' responses allows to obtain a high degree of favourableness and acquaintance capability of subjects to the humanoid. This preliminary research could be a first step to build a scale to assess the nature and efficacy of interaction with humanoid robots either in free spontaneous or in task driven contexts. Future research is needed to analyse whether item responses cluster within and across constructs, whether the constructs sub-dimensions could form a cognitive continuum or are instead two distinct cognitive scaffolds for interaction, whether all the features which are symbolized by the perceptual and social-cognitive clues reinforce mutually or may also come into conflict with one another. Caution is in order when the intent of constructing a scale is concerned, hence a further experimental probe of the balance, validity and face value meaning of the constructs is required. This exploratory descriptive analysis can provide a starting point for future research in the field.

References

- [1] S. Balistreri, G. Nishio, R. Sorbello, H. Ishiguro, Integrating built-in sensors of an android with sensors embedded in the environment for studying a more natural human-robot interaction, in: Lecture Notes in Computer Science, vol. 6934, Springer, Berlin/Heidelberg, 2011, pp. 432–437.
- [2] H. Ishiguro, S. Nishio, A. Chella, R. Sorbello, G. Balistreri, M. Giardina, C. Cali, Perceptual social dimensions of human-humanoid robot interaction, in: 12th International Conference on Intelligent Autonomous Systems on press in Advances in Intelligent and Soft Computing, 2012, ISSN: 1867–5662.
- [3] G. Balistreri, S. Nishio, R. Sorbello, A. Chella, H. Ishiguro, Natural human robot meta-communication through the integration of android's sensors with environment embedded sensors, in: Frontiers in Artificial Intelligence and Applications, vol. 233, IOS Press, 2011, pp. 26–37.
- [4] A. Chella, C. Lebiere, D. Noelle, A. Samsonovich, On a roadmap to biologically inspired cognitive agents, in: Frontiers in Artificial Intelligence and Applications, vol. 233, IOS Press, 2011, pp. 453–460.
- [5] T. Kanda, T. Miyashita, T. Osada, Y. Haikawa, H. Ishiguro, Analysis of humanoid appearances in human-robot interaction, IEEE Trans. Robot. 24 (2008) 725–735.
- [6] E. Oztop, D. Franklin, T. Chaminade, G. Cheng, Human-humanoid interaction: is a humanoid robot erceived as a human?, Internat. J. Humanoid Robot. 2 (2005) 537.
- [7] M. Poel, D. Heylen, A. Nijholt, M. Meulemans, A. Van Breemen, Gaze behaviour, believability, likability and the icat, AI Society 24 (2009) 61–73.
- [8] G. Metta, G. Sandini, D. Vernon, L. Natale, F. Nori, The icub humanoid robot: an open platform for research in embodied cognition, in: Proceedings of the 8th Workshop on Performance Metrics for Intelligent Systems, ACM, 2008, pp. 50–56.
- [9] H. Ishiguro, Android science-toward a new cross-interdisciplinary framework, Robot. Res. 28 (2007) 118–127.
- [10] T. Kanda, H. Ishiguro, T. Ono, M. Imai, R. Nakatsu, Development and evaluation of an interactive humanoid robot robovie, in: IEEE International Conference on Robotics and Automation, 2002. Proceedings. ICRA'02 vol 2, IEEE, 2002, pp. 1848–1855.
- [11] M. Shimada, T. Minato, S. Itakura, H. Ishiguro, Evaluation of android using unconscious recognition, in: 2006 6th IEEE-RAS International Conference on Humanoid Robots, IEEE, 2006, pp. 157–162.
- [12] M. Argyle, J. Dean, Eye-contact, distance and affiliation, Sociometry (1965) 289–304.
- [13] M. Argyle, R. Ingham, F. Alkema, M. McCallin, The different functions of gaze, Semiotica 7 (1973) 19–32.
- [14] N. Durlach, M. Slater, Presence in shared virtual environments and virtual togetherness, Presence: Teleoperators Virtual Environ. 9 (2000) 214–217.
- [15] S. Zhao, Toward a taxonomy of copresence, Presence: Teleoperators Virtual Environ. 12 (2003) 445–455.
- [16] M. Slater, A. Sadagic, M. Usoh, R. Schroeder, Small-group behavior in a virtual and real environment: a comparative study, Presence: Teleoperators Virtual Environ. 9 (2000) 37–51.
- [17] E. Goffman, Behavior in Public Places: Notes on the Social Organization of Gatherings, Press of Glencoe, 1966.
- [18] V. Gallese, Embodied simulation: from neurons to phenomenal experience, Phenomenol. Cognit. Sci. 4 (2005) 23–48.
- [19] V. Gallese, A. Goldman, Mirror neurons and the simulation theory of mindreading, Trends Cognit. Sci. 2 (1998) 493–501.
- [20] K. Dautenhahn, The art of designing socially intelligent agents: science, fiction, and the human in the loop, Appl. Artif. Intell. 12 (7–8) (1998) 573–617.
- [21] L. Cronbach, Coefficient alpha and the internal structure of tests, Psychometrika 16 (1951) 297–334.
- [22] J.C. Nunnally, Psychometric Theory 3E, Tata McGraw-Hill Education, 2010.
- [23] K. Dautenhahn, Socially intelligent robots: dimensions of human-robot interaction, Phil. Trans. R. Soc. B 362 (2007) 679–704.
- [24] M.L. Walters, K.L. Koay, D.S. Syrdal, K. Dautenhahn, R. Te Boekhorst, Preferences and perceptions of robot appearance and embodiment in human-robot interaction trials, Procs of New Frontiers in Human-Robot Interaction (2009).
- [25] C. Breazeal, Toward sociable robots, Robot. Auton. Syst. 42 (2003) 167-175.
- [26] M.S. Graziano, C.G. Gross, Spatial maps for the control of movement, Curr. Opin. Neurobiol. 8 (1998) 195–201.
- [27] C. Calì, Isomorphism and mirror neuron system, Gestalt Theory Int. Multidiscip. J. 29 (2007) 168-173.



Rosario Sorbello was born in Palermo, Italy on 6th March 1974. He received his Laurea degree in Computer Engineering and his Ph.D. degree in Computer Engineering from the University of Palermo, Italy in 1998 and 2001 respectively. He was a visiting Ph.D. student in 2000 in the Robotic Lab of Prof. Arkin. Currently, he is an Assistant Professor of Robotics at the University of Palermo. His research interests are in the field of Multi-Robot Teams. Autonomous Robots, Humanoid Robotics, and the human-humanoid interaction.



Giardina Marcello Emanuele was born in San Cataldo. Italy, on 7th June 1981. He received his Laurea degree in Computer Engineering from University of Palermo, Italy in 2011. Currently, He is a Ph.D. Student, in Computer Engineering at the University of Palermo. His research interests are in the field of Autonomous Robots, Humanoid Robotics, and the Human-Humanoid interaction.

Shuichi Nishio received his M.Sc. in Computer Science from Kyoto University in 1994, a D. Eng from Osaka

University in 2010, and is currently a Senior Researcher

at the Advanced Telecommunications Research Institute



Antonio Chella was born in Florence, Italy, on 4th March 1961. He received his Laurea degree in Electronic Engineering and his Ph.D. degree in Computer Engineering from the University of Palermo, Italy in 1988 and 1993 respectively. Currently, he is a Professor of Robotics at the University of Palermo. He is the head of the Robotics Lab of the University of Palermo since 2001. He has been the coordinator of the Course of Studies in Computer Engineering and the Head of Department of Computer Engineering. His research interests are in the field of robot consciousness, cognitive robotics and robot perception.



Carmelo Calí was born in Italy in 1972. He received a degree in Sciences and Humanities from the University of Bologna and a Ph.D. degree from the University of Palermo. Since 2004 he is an academic researcher at the University of Palermo. His research interests are in Cognitive Sciences, Philosophy of Mind and History of Psychology, and mainly focused on perception and abstract modelling for interdisciplinary research.







Hiroshi Ishiguro, a Professor of the Department of Systems Innovation in the Graduate School of Engineering Science, Osaka University and Fellow at the Intelligent Robotics and Communications Laboratories of the Advanced Telecommunications Research Institute (ATR), Japan, is now a Distinguished Professor of Osaka University (2013). After earning his Ph.D. in Systems Engineering from the Osaka University in 1991, he held posts as Associate Professor at the Department of Social Informatics, Kyoto University, Visiting Scholar at the University of California, Professor at the Department of Computer and

Communication Sciences, Wakayama University. His research involves distributed sensor systems, interactive robotics, and android science. He is a frequent lecturer on Robotics and Android Science worldwide

International (ATR) in Kyoto.