///VSP///

Advanced Robotics 23 (2009) 1951–1996

ADVANCED ROBOTICS
www.brill.nl/ar

Full paper

Effects of Embodiment and Gestures on Social Interaction in Drumming Games with a Humanoid Robot

Hatice Kose-Bagci *, **Ester Ferrari**, **Kerstin Dautenhahn**, **Dag Sverre Syrdal and Chrystopher L. Nehaniv**

Adaptive Systems Research Group, University of Hertfordshire, School of Computer Science, Hatfield, AL 10 9AB, UK

Received 9 June 2009; accepted 8 July 2009

Abstract

We present results from an empirical study investigating the effect of embodiment and minimal gestures in an interactive drumming game consisting of an autonomous child-sized humanoid robot (KASPAR) playing with child participants. In this study, each participant played three games with a humanoid robot that played a drum whilst simultaneously making (or not making) head gestures. The three games included the participant interacting with the real robot (physical embodiment condition), interacting with a hidden robot when only the sound of the robot is heard (disembodiment condition; note that the term 'disembodiment' is used in this paper specifically to refer to an experimental condition where a physical robot produces the sound cues, but is not visible to the participants), or interacting with a real-time image of the robot (virtual embodiment condition). We used a mixed design where repeated measures were used to evaluate embodiment effects and independent-groups measures were used to study the gestures effects. Data from the implementation of a human–robot interaction experiment with 66 children are presented, and statistically analyzed in terms of participants' subjective experiences and drumming performance of the human–robot pair. The subjective experiences showed significant differences for the different embodiment conditions when gestures were used in terms of enjoyment of the game, and perceived intelligence and appearance of the robot. The drumming performance also differed significantly within the embodiment conditions and the presence of gestures increased these differences significantly. The presence of a physical, embodied robot enabled more interaction, better drumming and turn-taking, as well as enjoyment of the interaction, especially when the robot used gestures.

Keywords

Human–robot interaction, embodiment, gestures, humanoid, drumming game, social interaction

* To whom correspondence should be addressed. E-mail: hatice.kose@gmail.com

 DOI:10.1163/016918609X12518783330360

1. Introduction

Social robots are being used widely in interaction games with human interaction partners, e.g., in the application areas of entertainment robotics [1–4], socially assistive robotics [5] and robot-assisted therapy [6–9]. Their physical appearance as well as behavior affect the participants and motivate them to take part in the interaction. The use of social cues such as gestures produced by the robots also has a great impact on the motivation of the human participants. In human–human interaction, gestures play an important role in communication, coordination and regulation of joint activities. In the related field of virtual agents, the beneficial effects of gestures and expressions used by virtual agents were shown both in short-term and long-term interactions, in maintaining user involvement with the tasks encouraged by the agent [10, 11].

Given that the term 'social robots' refers to robots that are designed to evoke meaningful social interaction with their users [12], social robots may not necessarily need a physical body to accomplish their goals [13], unlike other robotic systems where the task requires a physical form (e.g., object manipulation). Socially interactive robots are used in various kinds of applications, such as toys for entertainment, rehabilitation aids or educational tools. Often their primary functionality is not strictly related to physical interaction or manipulation, which implies that a physical body may not be required (a virtual embodiment could be sufficient). For example, in the field of rehabilitation robotics, which often focuses on physically assistive robots, non-contact socially assistive robots have been developed with the primary goal to motivate and monitor the user during the rehabilitation phase [14, 15]. Thus, this raises the question of whether physical embodiment is essential for successful interactions between a human and a social robot.

In the context of this paper we follow the notion of embodiment that has been defined as "that which establishes a basis for structural coupling by creating the potential for mutual perturbation between system and environment" [16, 17]. Note that this definition of embodiment does not necessarily require a system to possess a physical shape. According to Chrisley and Ziemke [18], embodiment can be classified in four different levels, from physical realization, where the system must simply be realised in some physical substrate, to organismal embodiment where the body must be alive (i.e., metabolize, reproduce, etc.).

Previous research has shown that physical embodiment has positive effects on the quality of interaction between social robots and humans. Lee *et al.* [19] conducted two experiments to investigate the effects of physical embodiment and tactile communication in human–agent interaction. They found that physical embodiment, as a bodily presence, played an important role in social interactions between human and social agents, although social robots were not particularly related to physical functions. Participants preferred interactions with physical social robots to interaction with virtual social robots. In the first experiment they found that physical embodiment positively impacted the agent's social presence, as well as the evaluation of the interaction with the agent, while a second experiment showed that physical embodiment with restricted tactile interaction caused insignificant or even negative effects in human–agent interaction.

Positive effects of physical embodiment have also been found by Bartneck [20] in his study with an emotional robot (eMuu). Specifically, he claimed that physical embodiment facilitated social interaction. In an empirical study, participants acquired a higher score in a negotiation game when they interacted with a robotic character than when they interacted with a character on a computer screen.

Tapus and Mataric [21] studied the effect of embodiment in a human–robot interaction (HRI) experiment with social robots playing music. Here a physical robot or a simulated computer animation played recorded songs and patients with cognitive impairments tried to distinguish between up to four songs. Future work needs to demonstrate whether the physical embodiment of the robot motivated the patients positively and helped them to improve their cognitive impairments.

Two experiments conducted by Wainer *et al.* [15, 22] add further support to the importance of physical embodiment on performance and impression of social interactions. Their results demonstrate that a physically situated robot is more appealing than a non-embodied robot, and it is also seen as more helpful, watchful and enjoyable when compared to a remotely tele-present robot and a simulated one. These results suggest that a physical robot may be more effective in assistive physiotherapy (the particular application area that the authors aimed at) than a disembodied one.

Conversely, a study conducted by Powers *et al.* [23] showed that an interaction with a 'collocated robot' (physical embodied robot) compared to a remote projected robot does not always lead to better results. They found that the projected robot had almost as much social influence as the collocated one, i.e., it was equally engaging, elicited equal disclosure, but may have had somewhat less influence (the participants did not rate the projected robot as highly when they evaluated its helpfulness, the usefulness of its advice and its effectiveness as a communicator).

Interestingly, all the above results are related to studies that have been conducted with adults. Children, even if they are the main target in the area of entertainment robotics, have not been involved extensively in such research comparing physical and virtual embodiment conditions. One of the few studies conducted with children participants has been carried out by Pereira *et al.* [2], involving 18 children in a gaming scenario against a robotic agent or a virtually embodied agent. In spite of the results of this study suggesting that embodiment has a positive effect on participant's enjoyment, there is still sparse evidence of the effect that embodiment has on children and further investigation is needed.

A social robot needs a set of social skills in order to successfully encourage a user's social behavior, which might require the ability to use social cues and gestures to motivate users to interact with it and keep them motivated to interact with the robot beyond the first few moments of 'novelty'. This is especially the case for assistive robotics [5]. A variety of robotic systems have been using social cues and gestures in order to encourage HRI, e.g., KISMET [24, 25], where the interaction itself was the primary goal. Different from this work, our own studies include an enjoyable task that will need to be achieved jointly by the human–robot pair. Therefore, we have chosen drumming as a test bed for our studies. Drumming is relatively straightforward to implement and test, and can be implemented technically without special actuators like fingers or special skills. Additionally, it is an easy and enjoyable task for the participants, who do not require any detailed information or skills. There are several approaches concerning drumming in HRI. For example, robotic percussionists play drums in collaboration with human interaction partners, where they use robotic arms that are specially designed to play drums [26, 27]. Similarly, humanoid drumming is used as a test bed for exploring synchronization [28].

Drum-mate is an interaction-based imitation game based on the autonomous drumming game of a human interaction partner and a humanoid robot [29, 30]. In Drum-mate studies, the humanoid robot KASPAR (Kinesics and Synchronization in Personal Assistant Robotics) plays drums autonomously with a human 'partner' (interactant), trying to imitate the rhythms produced by the human. However, the social interaction is not limited to the replication of drumming, but also involves studying the impact of non-verbal robot gestures that are meant to motivate the human. KASPAR produces head gestures from a limited repertoire and eye-blinking as it drums. KASPAR is a minimally expressive child-sized humanoid robot developed previously by our research group for its use in human–robot social interaction games (for more details of the robot, see Ref. [31]). In previous work, two studies with 24 adult participants each [29, 30] analyzed the interaction between participants and the humanoid robot in terms of imitation, turn-taking and the impact of non-verbal gestures as social cues [30]. Different computational probabilistic models were used to achieve turn-taking that is not deterministic, but emerging from the interaction between the human and the humanoid robot. The humanoid robot is no longer a passive 'follower', but can also play the 'leader' role in the game [29]. Also, different orders of the game conditions were tested and a significant effect of play time was found. The error rates in drumming and turn-taking significantly decrease as the human players play more games [29, 30].

The above-mentioned Drum-mate scenario (with adult participants) formed the basis for the current study which is a modified version of the Drum-mate game. It was tested with 66 primary school students, where different embodiment conditions, together with their relation with the head gestures, were studied. Each participant played three interactive drumming games with the humanoid robot. In each game the participant interacted either with the real robot (physical embodiment condition), with a real-time image of the robot (virtual embodiment condition) or with the hidden robot (disembodiment condition) — in this last condition only the sound of the robot is heard. Half of the children played games while the humanoid robot was simply drumming without making any head gestures; during the games with the other half of the children, KASPAR played its drum whilst simultaneously making gestures and waving its hand 'good-bye' at the end of the game.

Compared to the previous Drum-mate experiments with adults, several modifications were made in the current study to adapt the experiment to the child participants (e.g., simpler gestures were used, and the single game duration and the time between turns were decreased).

The paper is organized as follows. Section 2 presents the research questions and hypotheses. The experiment design and data collection are described in Section 3. In Section 4, the experimental results are described. Section 5 includes a brief conclusion of the experiment, lessons learned and presents ideas for future work.

2. Research Questions and Hypotheses

The goal of the study was to determine whether embodiment and gestures have an effect on how users perceive a social robot. In this experiment, we examined three levels of embodiment, each of which was used both with and without robotic gestures.

To study the effects of the embodiment, each child played a drumming game in the following three conditions:

- *K* The physical embodiment condition in which KASPAR sat on a table and played a drum in front of the child;
- *V* The virtual embodiment condition in which KASPAR's image was projected on the wall (keeping its real dimension in terms of size), while the robot played a drum behind an opaque barrier that separated it from the child;
- *D* The disembodiment condition (only sound), in which KASPAR and the child were in two different areas, and KASPAR played a drum behind an opaque barrier. Note that in this case participants were not able to see the robot, but they could actually hear when it was producing gestures (i.e., in the gesture condition the children could hear the robot's motors moving behind the screen).

Note that different areas had to be created for practical reasons, i.e., in order to allow each child to be tested in three experimental conditions and allowing quick changes between the experimental settings.

Based on the results of previous research (see Section 1), we expected that a social robot, in order to be able to engage in a playful interaction with a child, would require a certain degree of embodiment. Thus, we investigated the following hypotheses:

- H1 Children would evaluate a social robot and the interaction with it more positively when they played with an embodied robot (conditions K and V), than when they interact with a disembodied robot (condition D); and comparing the physical and virtual embodiment conditions, they will evaluate the physically embodied social robot (K) more positively. Moreover, we expected that the presence of gesture would increase the difference in how the children evaluate the embodiment conditions. Specifically, we expected the children would evaluate the case K (physical embodiment) more positively than the other cases (V and D) when gestures were used.

H2 The error rates in drumming and turn-taking would decrease when the children played with an embodied robot (conditions K and V) compared when they interact with a disembodied robot (condition D); and comparing the physical and virtual embodiment conditions, the error rates in the physically embodied robot (K) condition will be lower than the virtually embodied robot (V). The differences between the drumming and turn-taking performances when the children play with different conditions (K, V or D) are expected to be higher when the gestures were introduced.

H3 As the play time increases, the error rates would decrease. In other words, the more the child plays, the better her/his drumming and turn-taking would be. Therefore, we expect their drumming and turn-taking to improve over time and, consequently, we expect a better performance (lower error) in the third game played than the first game played. The differences between the drumming and turn-taking performances between first and third games are expected to be higher when the gestures were introduced.

To study the effects of gestures, the robot played half of the games while making a gesture ('gesture' condition) and half of the games without making any gesture ('no-gesture' condition).

In addition to the above hypotheses, we will also consider the effects of gender. Previous works show gender differences have an impact on the subjective and objective evaluation of human participants. Kose-Bagci *et al.* showed that females and males evaluate the robot and the interaction games differently, and their drumming and turn-taking performances differ significantly [30]. Here the males were more 'task oriented', whereas females tended to value interactional aspects of the scenario. However, in Ref. [32] it was found that the males like to see the robot as more 'human-like' and achieve a social facilitation, while females saw it 'machine-like'. Gender differences were also revealed in HRI experiments [33–35]. This suggests the possibility of gender as a confounding variable in this experiment, which will be examined in the data analysis below. However, gender issues do not play a major role in our research goals (and for this reason our sample is not gender balanced).

3. Experiment

3.1. Participants and Sample

Sixty-six participants in the age range of 9–10 years took part in the study. All participants were primary school students from six schools in Hertfordshire, UK. Gender was not balanced in the sample; the majority of the children were female ($n = 39$, Table 1). None of the children had interacted with the robot KASPAR prior to the experiment. Most of them were used to playing computer games (Table 2)

Table 1.
Distribution of the sample's gender

	N	%
Male	27	41
Female	39	59
	66	100

Table 2.
Distribution of the sample's familiarity with robots

	N	%
I've played with a robot before	19	29
I've never played with a robot	47	71
	66	100

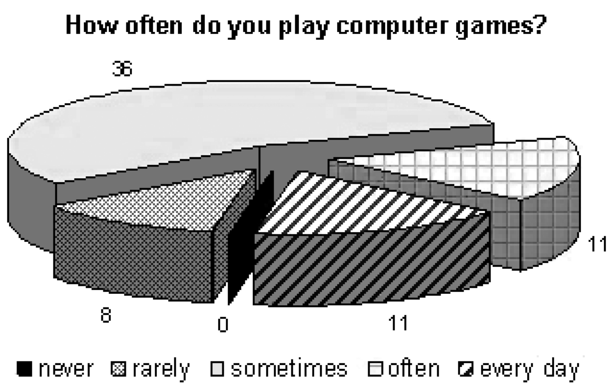

Figure 1. Distribution of how often the children play computer games.

and they were generally unfamiliar with robots (Fig. 1). Prior to the experiment the children's parents had consented for their children to take part in the study, and allowed us to obtain photos and video recordings of the experiments for scientific purposes.

3.2. Design

To test the hypotheses, a 3 (embodiment) × 2 (gestures) mixed design was used. The experiment consisted of drumming games with a humanoid robot in three different embodiment conditions (physical embodiment, virtual embodiment and disembodiment) and two gesture conditions (with and without gestures).

Each participant played an equivalent drumming game with the robot in the three different embodiment conditions, each of which took 2 min (making embodiment a repeated measures variable). Half of the children were randomly assigned to the gesture condition while the remaining half was assigned to the no-gesture condition (making gestures an independent groups variable).

In the no-gesture condition, KASPAR only played the drum without making any gestures, while in the gesture condition, the robot made simple head gestures, e.g., nodding and moving the head from side to side during its drumming session. These head gestures were played in a fixed sequence to encourage the participants to believe that they were executed on purpose and not at random [30]. KASPAR smiled when it started drumming and if it did not detect the child's drumming when expected (e.g., if the drumming beats were too light, too fast or if the child did not play), then it blinked and expressed a neutral smile. At the end of the game KASPAR waved its hand with a 'good-bye' gesture to notify the end of the game. Note that in the disembodiment condition participants did not see the gestures, but they could hear them (i.e., they could hear the robot's motors moving behind the screen) and the production of the gestures also slightly influenced the robot's behavior timing. Thus, for completeness purposes we also included the condition where the disembodied robot used gestures.

A repeated measures design has the advantage that individual differences between participants are removed as a potential confounding variable, but a drawback related to the order effects. As we assumed that the order in which the three embodiment conditions were presented could influence the participant's opinion and behavior, their order was counterbalanced and all six possible presentation orders were used. This was essential to account for possible fatigue, habituation or learning effects.

3.3. Experimental Setup

The experiment was conducted during the event called 'Take Part In The Future And FearNot!' [36], hosted by the University of Hertfordshire, School of Computer Science and School of Education in May 2008, where 8- to 12-year-old children had the opportunity to interact with robots and trial anti-bullying software. In addition, in a different large room, the children were also able to interact with a number of humanoid and non-humanoid robots that are used in our research group. While screens were used to separate the experimental area from the other robotics demonstrations, the experimental setting was challenging as it was not an easily controlled laboratory environment.

Moreover, in conducting the experiment the enjoyment of the activities for the participants was taken into account and children were encouraged to experience an interaction with a social robot in an enjoyable manner. While this setting might have made it easier for children to express a more positive opinion than they would in a different setting (all the variables' means were quite high), such a setup provided an enjoyable and relaxing environment that creates situations more similar to those where children's play naturally occurs. It has been argued by Sabanovic *et al.* [37] that 'Interactions with robots in the laboratory, under the watchful eye and expert guidance of the robot's designers, do not provide insights into the aspects of human robot interaction that emerge in the less structured real-world social settings in which they are meant to function. It is, therefore, necessary to evaluate human–robot interactions as socio-culturally constituted activities outside the laboratory, or "in the wild" ' (p. 576). Studying social and enjoyable games that children play with a robot requires a suitable setting, not one where the children may be under the impression that they are being evaluated or monitored (similar to an examination). Thus, we had to find a trade-off between such an enjoyable setting and the need to control the experiments. The solution we decided on is to have an experimental setup on University premises, but to situate it in the context of an enjoyable activity for the children. Also, our sample was an 'opportunity sample' — we could not freely select the children we worked with, which meant that we could not control all the variables (e.g., gender). Thus, our research approach is similar to other HRI experiments in public places, museums, shopping malls, school environments, etc. (e.g., Refs [37–39]).

We designed a separate experimental area in the room where the robotics demonstrations took place. Two almost identical cubicles isolated from the rest of the room by tall screens were used to carry out the study (Fig. 2). In the remainder of the room other robotic activities took place at the same time with other children. In the first cubicle (area 1), the robot KASPAR was seated on a table with a toy drum on its lap. A chair was placed in front of the robot where the participant was seated (Fig. 3). In the second cubicle (area 2), there was a table and a chair for the participant. In the virtual embodiment condition, a real-time image of KASPAR (the same size as the physical robot) was projected on the wall in front of the seated child participant. In the disembodiment condition, the projector was switched off and the child participant could just hear the drumming sound of the robot hidden behind a screen (Fig. 4). To study the effect of the different embodiment conditions, all the others features of the setup in the cubicles were kept the same.

During the experiments, a drum (with a microphone attached) and a stick to hit the drum were made available to the children to make it easier for the robot to recognize the drumming sound through audio analysis. Although some of the children used the stick to hit on the top of the drum as we suggested, others preferred to use both their hands or to hit the tambourine-style bells around the drum's sides, which increased their enjoyment and involvement in the game, but negatively effects the audio analysis. Note that since the interaction was meant to be enjoyable and playful, we did not insist on the children using the drum stick.

Participants were instructed that they could play drumming games with the robot KASPAR. Simple general instructions about the game were given (e.g., hit the drum strongly so that KASPAR can 'hear' you better).

The children entered the experimental area in groups of three. On arrival, one of the experimenters gave them a 30-s demo of the first condition they were going to play. That was done simply to show them what the game was about and to familiarize them with the setup of the cubicle (e.g., if the first condition was 'vir-

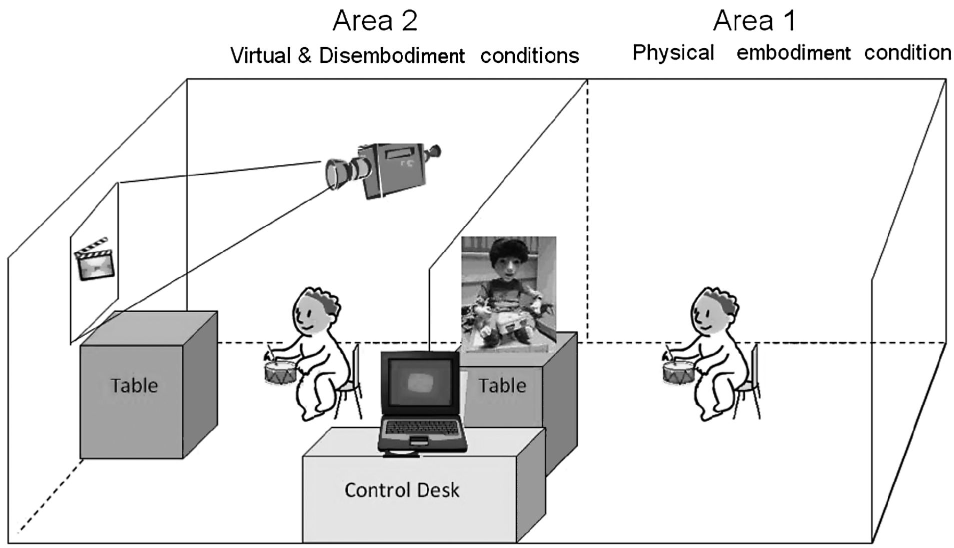

Figure 2. Setup of the experiment. The experiment took place in two identically designed cubicles. In the right cubicle (area 1), the game with the physical embodied condition was played, with KASPAR sitting on a table in front of the child. The left cubicle (area 2) was used to play the disembodiment and the virtual embodiment conditions. In the virtual embodiment condition, a real-time image (the same size as the physical robot) of the drumming robot was projected on the wall just above the desk at a height comparable to the physical embodied condition. In the disembodiment condition, the projector was switched off and only the sound of the robot hidden on the other side of the screen was presented. All the other features in the cubicles were identical. In all three conditions the child was playing alone in the cubicle. The control desk with the laptop controlling the robot was equidistant to both cubicles and the experimenter could be seen from both cubicles. Several cameras were located in the room to record the experiment.

tual embodiment', then the demo was carried out in area 2 and the children would see the experimenter playing with the projected image of KASPAR — not with the physical robot itself).

After the demo, only one child was asked to remain in the cubicle — the other two children were taken by the other experimenter outside the cubicle to wait for their turn. Each child individually interacted with the physically embodied (K), virtually embodied (V) and disembodied robot (D), in one of the two gesture conditions.

During the games, the experimenter who was located in the experimental area kept silent and did not interfere with the child's or the robot's performance. The children were given very basic information about the difference between the games (e.g., 'Now you will play with the projected image of KASPAR'; 'Now you will hear KASPAR's drumming sound, but you cannot see KASPAR itself').

The sessions were videotaped and the video recordings were used as a source of behavioral data (described in detail in the following subsections).

In the current study, KASPAR acted totally autonomously. Thus, the experimenter was always in view of the participant, especially when the physical robot

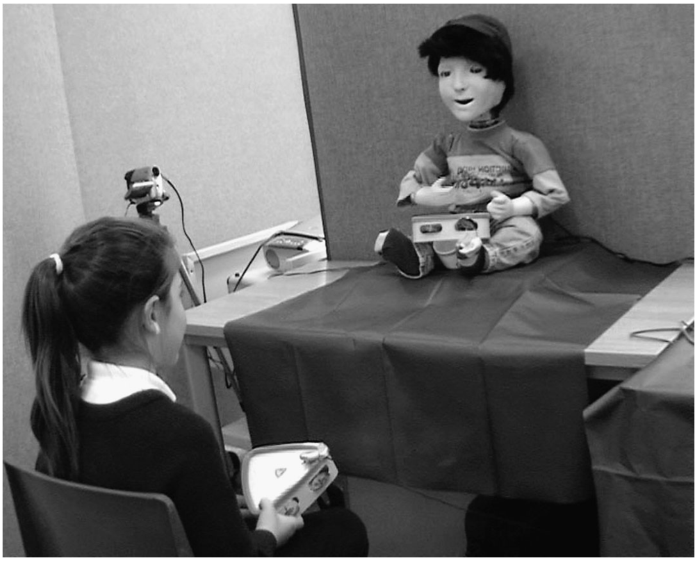

Figure 3. Screen shot from the experiment showing a person playing a drumming game with the physically embodied robot.

Figure 4. Screen shot from the experiment showing a person playing a drumming game with the hidden robot (disembodiment condition).

was not visible. Participants may otherwise have believed that the performer of the drumming was not KASPAR, but the experimenter (i.e., a hidden puppeteer in the Wizard-of-Oz technique — a widely used technique in human-computer interaction and HRI research where a human, who is unknown to the participants, is controlling the behavior of the system; e.g., Ref. [40]).

3.4. KASPAR

The study was carried out with the humanoid robot called KASPAR. KASPAR is a child-sized humanoid robot that was designed and built by the members of the Adaptive Systems Research Group at the University of Hertfordshire to study HRIs with a minimal set of expressive robot features. This humanoid robot has been used in a variety of projects, e.g., in research to mediate interaction for people with and without special needs [41, 42].

KASPAR has 8 d.o.f. in the head and neck, and 6 d.o.f. in the arms. Its width is 30 cm, depth is 35 cm and height is 45 cm, and its shape is modeled after a 2- to 3-year-old child. The face is a silicon-rubber mask, which is supported by an aluminum frame. It has eyelids capable of blinking, and a mouth capable of opening and smiling (a detailed description can be found in Ref. [31]).

3.5. Interaction Game Implementation

In this work, as in the previous study [30], the participant played a rhythm that KASPAR tried to replicate in a simple form of imitation (mirroring). KASPAR had two modes — listening and playing. It recorded and analyzed the human's rhythm in the listening mode and it played the rhythm back by hitting the drum positioned on its lap in the playing mode. Then the participant played again. This (deterministic) turn-taking in this game continued for a fixed time duration (2 min for the current work). Due to its limited motor skills, KASPAR did not imitate the strength of the beats, but only the number of beats and durations between beats. For beats beyond its motor skills, it used instead minimum values allowed by its capabilities: KASPAR needed at least 0.3 s between beats to get its joints 'ready', so that, even if the human played faster, KASPAR's imitations would still require minimum durations of at least 0.3 s between beats. It also needed to wait for a few seconds before playing any rhythm in order to get its joints into the correct reference positions.

3.6. Software Features

The implementation of robot perception and motor control used the YARP environment [43]. YARP is an open-source framework used in the project RobotCub that supports distributed computation and emphasizes robot control as well as efficiency. It enables the development of software for robots without considering a specific hardware or software environment. Portaudio [44] software was used to grab audio from the audio device within the YARP framework.

The acoustic sound waves recorded by the sound grabber module were converted to digital music samples, which allows mathematical computations and sample-based techniques to be used on them. To detect the patterns of a sound wave, a filter-based method is used, based on the work of Ref. [45] that was originally used to detect visual patterns.

3.7. Measures

During the experiment several sources were used to collect data. These sources included asking the children to complete a questionnaire related to the trials, recording the sessions by video cameras for later video analysis and collecting the behavioral (drumming) data of each robot–child pair.

3.7.1. Questionnaire

A paper-based questionnaire was used to collect data relevant in investigating the differences in the embodiment conditions (H1) and the gesture effect (H2). The questionnaire is available from E. F. request. It was comprised of three sections. The first section gathered general information about the child, their experience with robots and video games, and gave instructions on how to complete the questionnaire. The second section consisted of 15 closed-ended questions (repeated for the three embodiment conditions) and the third section consisted of two open-ended questions.

Considering the sample population and time limitations, the questionnaire was kept as simple and short as possible. It was pilot tested with one participant of a similar age group (a 9-year-old child). The pilot test confirmed that both the length of the questionnaire and the questionnaire's administration time were acceptable (below 6 min). In order to develop the final version of the questionnaire, a few minor language changes were made.

After each trial condition, the child was asked to answer the 15 closed-ended questions in order to express their opinions about the game they just played (the same questions were presented after each of the three embodiment trial conditions).

The closed-ended questions were used to evaluate the robot in terms of: enjoyment, social attraction, involvement, performance, general appearance and intelligence. As the researchers were interested in the children's feelings and opinions about the interaction with the robot, a five-point Likert scale (respondent shows the amount of agreement/disagreement with a given statement) and a semantic differential scale (a scale inscribed between two bipolar words; children select the point that most represents the direction and intensity of their feelings) were used.

Note that before conducting the data analysis, three items were removed from the questionnaire because some of the children, while they were completing the questionnaire, showed difficulties in understanding their meaning (one was a negatively phrased item).

To measure the level of enjoyment during the interaction with the robot, two questions on a five-point Likert scale with a central anchor were used ('Did you enjoy playing with this robot?' and 'Did you find it interesting?') (Cronbach's $\alpha = 0.78$). Social attraction toward the robot was measured by a modified version of McCroskey and McCain's Interpersonal Attraction Scale [46]; children were asked to indicate their level of agreement to the following statements: 'I would like to be friend with this robot' and 'I would like to spend more time with this robot' (Cronbach's $\alpha = 0.90$). Involvement in the game was measured by the level of agreement to the following statements: ‘I paid attention to the robot’ and ‘I felt that the robot involved me in the game’ (Cronbach’s $\alpha = 0.84$). The robot’s perceived intelligence was measured by participants’ levels of agreement on the statement: ‘I think this robot is intelligent’ (Cronbach’s $\alpha = 0.60$). The level of agreement was always measured using the same five-point response scales with a central anchor. Two questions concerning the robot’s performance were asked using a five-point semantic differential scale: lazy/energetic, and bad drummer/good drummer (Cronbach’s $\alpha = 0.79$). Three questions concerning the general appearance of the robot were asked using a five-point semantic differential scale: unpleasant/pleasant, not friendly/friendly and machine like/human like (Cronbach’s $\alpha = 0.78$).

Once the children had completed the items related to the third game, they were asked to complete the last section of the questionnaire. In this part they judged the overall experience by deciding which of the three games they liked the best and which they liked the least, as well as writing down the reasons behind that decision.

During the study, the questionnaire administration was performed in a dedicated area separate from playing areas 1 and 2, where one of the experimenters was on hand to help the child complete the questionnaire if needed.

3.7.2. Behavioral Data

The experiments were recorded by two different cameras positioned at different parts of the experimental area (one facing the child and the other facing the robot), during each single game. The video recordings were later analyzed manually to detect the performance of the children’s behavioral data (e.g., the number of drum beats played by the children and number of turns taken by the children at each game). This data was then compared with the behavioral data recorded by the robot itself (see below). Also, video recordings are helpful as they give valuable clues about the likes/dislikes of the children. They are also used to support the evaluation of the questionnaires.

Behavioral data that belonged to the robot and human participants were collected during the trials by the robot using its internal (joints) and external sensors (microphones). KASPAR records its performance (e.g., the number of drum beats played by KASPAR and number of turns taken by the robot during each game), as well as those of the children (e.g., the duration of time between each drum beat of the children) to imitate their performance within its physical limitations. In the next section these recordings will be described and analyzed in detail.

The behavioral data includes several parameters related to the children’s and to KASPAR’s drumming, i.e., the number of turns in a specific game, total, average and maximum number of drum beats performed by human participants and KASPAR per turn, and the drumming and turn-taking errors. The drumming error is the difference between KASPAR’s actual drumming, i.e., the number of beats KASPAR plays in a particular turn, and the number of beats the child plays. Likewise, the turn-taking error is based on the difference between KASPAR’s and the children’s turn-taking. Thus, the drumming and turn-taking errors reflect the discrepancy between human and robot drumming performance in this imitation game.

Although the robot's performance is the same under all conditions, the child's response to the robot's play differs. This affects the robot's detection and imitation of the child's drumming and, thus, influences the robot's performance in its response. There are many different reasons for the robot's erroneous detection of the child's performance (number of drumming beats and turns) caused by the children. For example, they will beat the drum while the robot is drumming and not listening (improper synchronization), so their beating will not be considered by the robot. Likewise, they may beat very fast or very light, which will not be detected by the robot, or they will use the bells of the drum, resulting in the robot detecting more than one beat.

In Section 4, several 'error' and performance measures based on the behavioral data are used to analyze the differences between different conditions. The error and performance measures are either presented per game or per turn. A game comprises the whole interaction occurring within one embodiment condition in a limited time period (2 min as specified in the current work) including several turns. The term non-zero turns defines the turns where at least one drum beat is played. For clarity all the zero turns (i.e., turns where no drum beat is played/or detected) were removed from the data for the following analysis. The term *Diffsum* (1) stands for the difference of the total sum of beats between participant and robot per game. The maximum number of beats per game shows the maximum number of drum beats played in a single turn per game. As shown in (2), *Errorsum* is *Diffsum* per number of non-zero turns (maximum of human and KASPAR). The term non-perfect turns is used for the number of non-zero turns where the number of drum beats in both the human's and KASPAR's turns do not match. If the number of non-zero turns of both do not match then the difference is also counted as a non-perfect turn. However, due to errors in observations and differences in children's play rhythms, this measure can be erroneous, giving a higher error rate than the real case, so we also take *Errorsum* and other performance measures into consideration. The term *Errorturn* defines the number of non-perfect turns per number of non-zero turns (maximum of human and KASPAR) (3).

$$\mathit{Diffsum} = \sum \text{Beats}_{\text{Human}} - \sum \text{Beats}_{\text{KASPAR}} \tag{1}$$

$$\mathit{Errorsum} = \frac{\mathit{Diffsum}}{\max(\text{non-zero\_turns}_{\text{Human}}, \text{non-zero\_turns}_{\text{KASPAR}})} \tag{2}$$

$$\mathit{Errorturn}\colon \frac{\text{non-perfect\_turns}}{\max(\text{non-zero\_turns}_{\text{Human}}, \text{non-zero\_turns}_{\text{KASPAR}})}. \tag{3}$$

To evaluate the success of a performance, the error rates, especially *Errorsum* and *Errorturn*, are taken into consideration. The lower the *Errorsum* and *Diffsum*, the better the drumming. Similarly, lower *Errorturn*, difference of non-zero turns or number of non-perfect turns values indicate better turn-taking. Ideally, *Errorsum* and *Errorturn* should be smaller than 1, and as close to 0 as much as possible. Other criteria, e.g., number of non-zero turns, average or maximum number of beats played per turn and number of beats played per game, can differ according to different conditions in the game or different features of the participants. For example, a higher average number of beats per turn might indicate more involvement of the human participant in that particular game compared to the other conditions, even though this might increase the errors in the game, due to the technical limitations of KASPAR's audio capture.

4. Results and Discussion

As mentioned above, the present study utilized a 3 (embodiment) $\times$ 2 (gestures) mixed design to evaluate embodiment and gesture effects. Data of the 66 children were analyzed using SPSS software (version 16 for Windows) and results are reported below. Detailed information about the descriptive data related to the questionnaires and behavioural data is listed in Appendix A.

In the following subsections the letter N stands for 'no-gesture' condition and the letter G stands for 'gesture' condition, i.e., they indicate whether the gestures of KASPAR were used in that particular game or not. Likewise, as explained above for the embodiment conditions, K stands for the game where the human participant played with the physical robot KASPAR, V is the virtual embodiment condition where the human played with the projected image of KASPAR and D is the disembodiment condition where the participant cannot see the robot, but can only hear the sound of the hidden KASPAR.

There are two main sources of error — the differences in KASPAR's and the child's drumming, and the differences in their turn-taking. The video recordings and the data recorded by the robot itself were analyzed to obtain these error and performance measures, which are very useful to study the behaviors of the robot and the children, and detect some significant differences between various conditions that are hard to detect from the questionnaire data only.

4.1. Does Embodiment Matter?

Data collected from the questionnaires were analyzed to investigate differences in children's opinions about their interaction with the robot in the different embodiment conditions.

In our research we predicted that the children would evaluate the interaction with an embodied robot more positively than the interaction with a disembodied one and that the interaction with the virtual robot would be less positive than the interaction with the physical embodied robot (as described in H1).

Hypothesis H1 is partially supported by the answers that the children gave to the overall experience. Two questions at the end of the questionnaire focused on collecting information about the game that they liked the most (Fig. 5) and the game that they liked the least (Fig. 6). As shown in Table 3, almost all the participants ($n = 55$; 83.33%) preferred to play with the embodied robot (conditions V or K), rather than with the disembodied one. Likewise, more than half of the children preferred the game in the physical embodiment condition ($n = 38$; 57.57%), compared

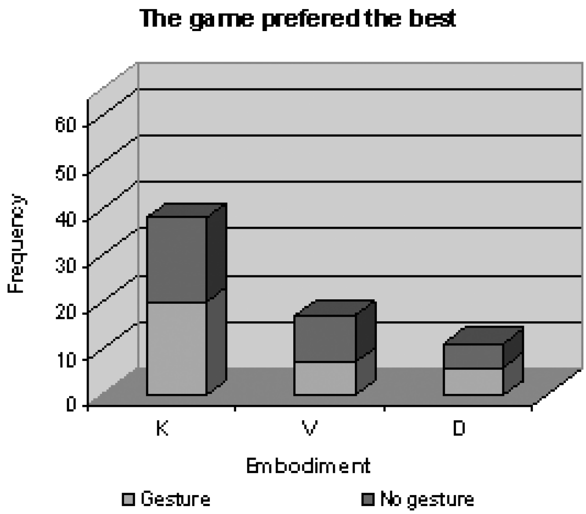

Figure 5. Most liked game.

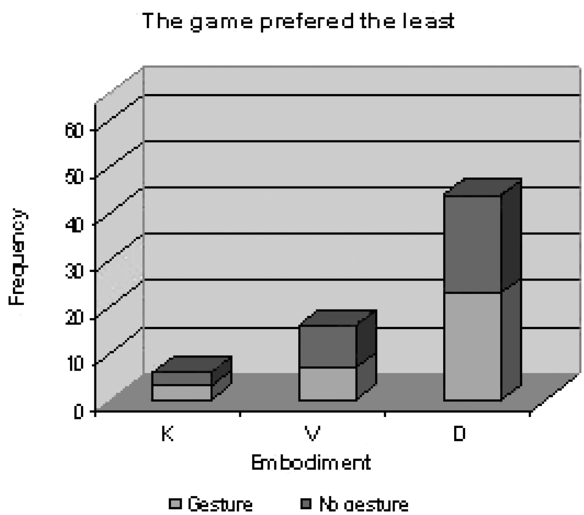

Figure 6. Least liked game.

to the game in the virtual embodiment condition (25.76%). In addition, they liked least the game played in the disembodiment condition ($n = 44$, 66.6%). However, no significant mean difference in the embodiment condition related to evaluation of enjoyment, social attraction, involvement, performance, general appearance and intelligence was found.

It is nevertheless interesting to note the trend that the data shows (Fig. 7). The interaction with the physically embodied robot was generally more appreciated than the interactions in the other two conditions. Indeed, in almost all the variables (e.g., robot's appearance, social attraction, involvement and intelligence) the

Table 3.
Game preferences

Embodiment	Game liked the best				Game liked the least			
	Frequency	Frequency in G condition	Frequency in N condition	%	Frequency	Frequency in G condition	Frequency in N condition	%
K	38	20	18	57.57	6	3	3	9.09
V	17	7	10	25.76	16	7	9	24.24
D	11	6	5	16.67	44	23	21	66.67
Total	66			100	66			100

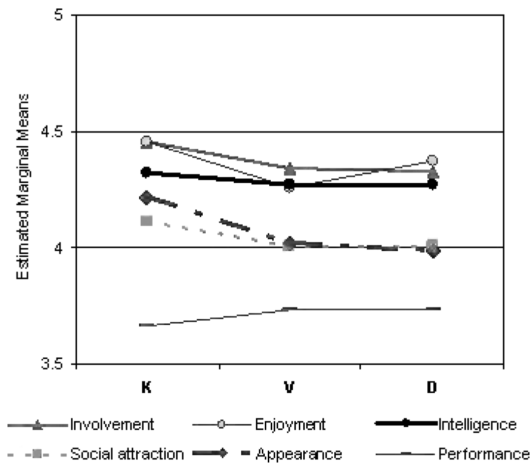

Figure 7. Effect of embodiment in children's scores of robot appearance, social attraction, involvement, intelligence, performance and enjoyment of the interaction.

children gave the physically embodied robot the highest score, the disembodied robot the lowest score and the virtual embodied robot a score in between the two (generally close to the one assigned to the disembodiment condition). On the contrary, a different result appeared for the level of enjoyment. It is interesting to notice that the virtual embodiment condition received the lowest score — lower than the disembodied condition. Similarly, another result that does not follow the previous highlighted trend is the one related to the robot's performance. The disembodiment condition gained the highest score in robot performance, while the physically embodied robot received the lowest score. A possible explanation for this unforeseen result is that children's attention, while interacting with the disembodied robot, is not diverted from the primary task, so children were focused only on the drumming game and, thus, evaluated the drumming performance differently.

However, consistent with hypothesis H1, results show that when gestures were used, the participants tended to evaluate the physically embodied robot more pos-

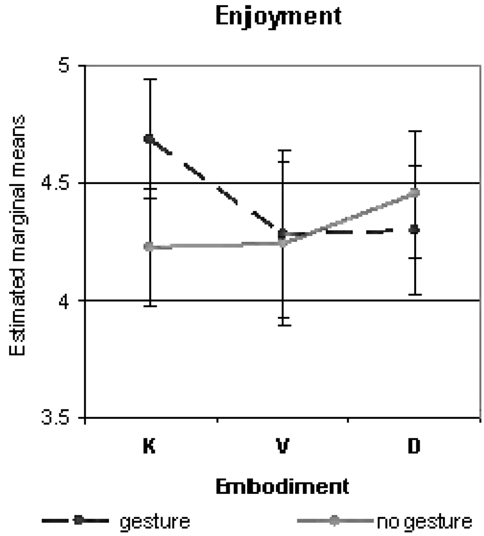

Figure 8. Combined effect of embodiment and gesture in children's scores of the enjoyment of the interaction.

itively than in the other two conditions. In terms of enjoyment, there was an interaction effect between gesture and embodiment ($F(2,126) = 4.29$, $P < 0.016$, $\eta^2 = 0.064$). This effect is described in Fig. 8, which suggests that for the gesture condition, participants tended to evaluate the physical condition more favorably than other conditions — an effect which is not evident for the no-gesture condition.

Likewise, there was an interaction effect between gestures and embodiment in terms of perceived intelligence ($F(2,126) = 3.24$, $P < 0.042$, $\eta^2 = 0.049$). This effect is described in Fig. 9, which suggests that the physical robot is rated as more intelligent in the gesture condition while the opposite is true for the virtual embodiment and the disembodied condition. It may be that the time the robot spent in making gesture movements affected negatively the intelligence attributed to it when it was not physically present (V and D condition), while it could have had a positive effect for the robot in the K condition.

It is also interesting to notice that the disembodied robot (D) in the no-gesture condition and the physically embodied robot (K) in the gesture condition received a similar score. This result might be related to the number of drum beats played per turn (see Section 4.2).

Moreover, a significant interaction effect between gestures and embodiment was found in terms of robot appearance ($F(2,128) = 4.92$, $P < 0.009$, $\eta^2 = 0.071$). Note, as mentioned, children neither saw the robot's appearance nor its gestures during the disembodiment condition, so we refrain from discussing in more detail any results concerning the disembodiment condition with respect to appearance evaluation or effects of gestures. Figure 10 suggests that this effect caused the children to evaluate the robot's appearance in the virtual condition more positively when ges-

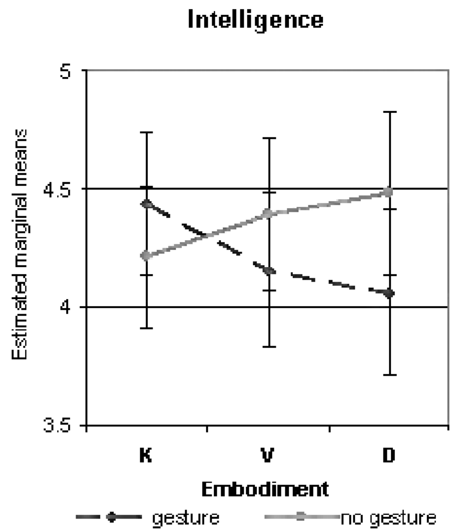

Figure 9. Combined effect of embodiment and gesture in robot's intelligence.

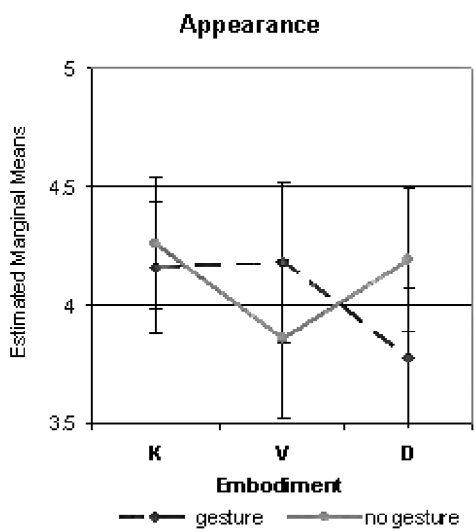

Figure 10. Combined effect of embodiment and gesture in robot appearance.

tures were used than when they were not used. Different from the previous trend, children assigned a higher score to robot appearance when the physically embodied robot was drumming in the no-gesture condition than in the gesture condition. It might be the case that the gestures made by the robot were appropriate for a projected image of the robot, but not smooth enough for a robot sat in front of the participants.

No significant interaction effects between the gesture conditions and embodiment conditions concerning social attraction, performance and involvement were found.

4.2. Effect of Embodiment and Gestures on Behavioral Data (H2)

The drumming and turn-taking performances differ for different embodiment conditions based on the analysis of behavioral data including error rates on the drumming and turn-taking, and several other parameters, i.e., total amount of drumming per game and maximum number of drum beats per turn, as stated in Section 3.7.2.

The game types are compared in detail in Table 4 (differences between human and KASPAR's perspectives), Table 5 (human's perspective) and Table 6 (KASPAR's perspective). As shown in Fig. 11, there is a highly significant difference between the physically embodied condition *K* and the other conditions

Table 4.
Observed differences between child–robot drumming behaviors according to the embodiment condition

Game type	Difference of sum of beats	Difference of non-zero turns	No. of non-perfect turns	Error in sum of beats	Error in number of turns
K	12.53 ± 11.0	1.86 ± 2.00	14.65 ± 4.1	0.62 ± 0.51	0.74 ± 0.13
V	18.38 ± 13.2	2.65 ± 3.40	14.98 ± 4.1	0.94 ± 0.66	0.78 ± 0.14
D	18.20 ± 16.5	2.47 ± 2.39	15.45 ± 4.0	0.90 ± 0.80	0.78 ± 0.14

Table 5.
Observed human drumming behavior according to the embodiment condition

Game type	Sum of beats	No. of non-zero turns	Maximum no. of beats	Average no. of beats/turn
K	60.06 ± 4.53	19.45 ± 4.02	12	3.18 ± 0.89
V	67.73 ± 17.72	18.70 ± 4.87	25	3.89 ± 1.73
D	67.61 ± 18.70	19.11 ± 4.18	13	3.64 ± 1.08

Table 6.
Observed KASPAR's drumming behavior according to the embodiment condition

Game type	Sum of beats	No. of non-zero turns	Maximum no. of beats	Average no. of beats/turn
K	51.86 ± 15.22	18.11 ± 3.57	14	2.93 ± 0.97
V	54.68 ± 17.81	17.59 ± 4.22	23	3.31 ± 1.73
D	53.35 ± 17.51	18.21 ± 3.52	13	2.98 ± 1.10

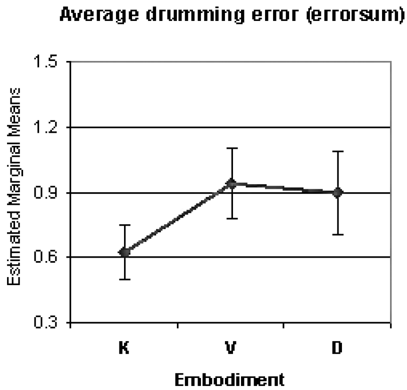

Figure 11. Average errors in sum of beats according to the embodiment condition.

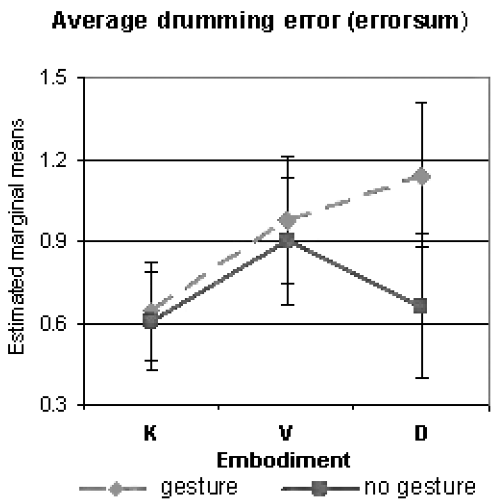

Figure 12. Effect of gestures on the average errors in the sum of beats according to the embodiment condition.

in terms of the average error in the sum of beats (*Errorsum*) ($F(2,126) = 6.653$, $P < 0.002$, $\eta^2 = 0.094$; Bonferroni adjusted *post-hoc* tests indicate that the mean difference between K and V is 0.315, $P = 0.047$; the mean difference between K and D is 0.276, $P = 0.001$). The error rate increases with the absence of physical embodiment, but not between disembodiment and virtual embodiment conditions. There is a significant interaction effect between the gesture condition and embodiment condition ($F(2,126) = 3.379$, $P < 0.037$, $\eta^2 = 0.050$). This interaction effect, presented in Fig. 12, suggests that the differences found between condition K and the others is more pronounced for the gesture condition. There is no significant main effect between embodiment conditions for turn-taking error (see Fig. 13), but a significant interaction effect ($F(2,126) = 8.214$, $P < 0.000$, $\eta^2 = 0.114$) presented in Fig. 14 suggests that gestures decrease the likelihood of such errors in

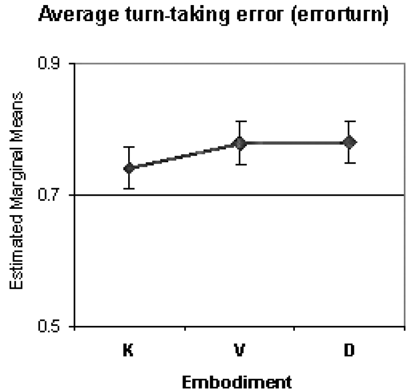

Figure 13. Average errors in number of turns according to the embodiment condition.

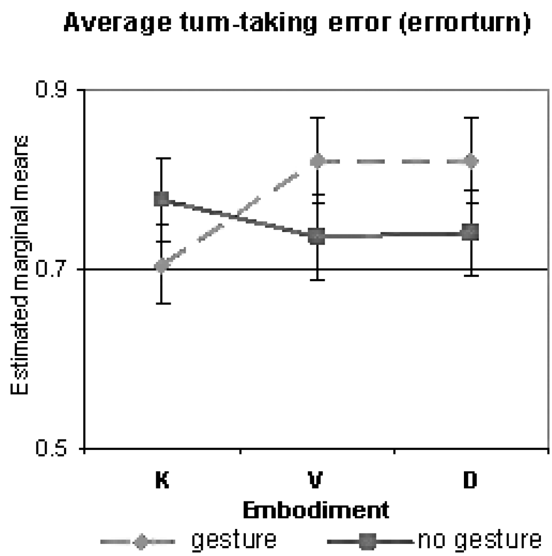

Figure 14. Effect of gestures on the average errors in number of turns according to the embodiment condition.

the physically embodied condition, while having the opposite effect on the virtual and disembodied condition. These results support hypothesis H2 and suggest that the physically embodied robot helps the child to understand the game and the robot better, and improve the drumming and turn-taking performances.

The difference in the sum of drum beats is also significantly higher in the V and D conditions than in condition K (main effect: $F(2,126) = 6.096$, $P < 0.003$, $\eta^2 = 0.087$; Bonferroni adjusted *post-hoc* tests indicate a mean difference between K and V of 8.939, $P = 0.0.001$; the mean difference between K and D is 7.591, $P = 0.0001$), while there are no significant differences between the V and D conditions (Fig. 15). An interaction effect between gesture and embodiment conditions,

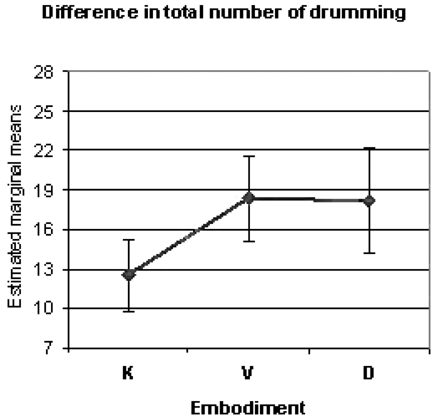

Figure 15. Difference of total number of drum beats of children and KASPAR according to the embodiment condition.

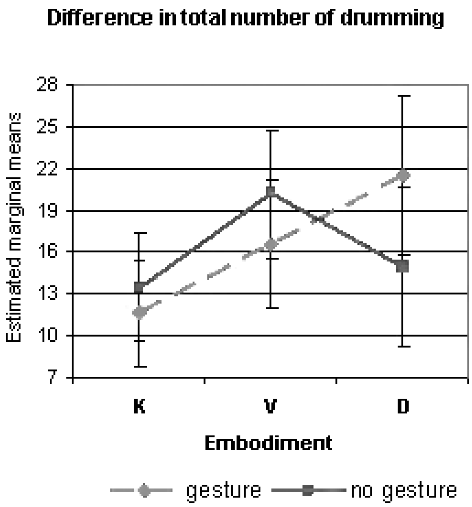

Figure 16. Effect of gestures on the difference of total number of drum beats of children and KASPAR according to the embodiment condition.

presented in Fig. 16, suggests that this effect is more strongly pronounced for the gesture condition ($F(2,126) = 4.031$, $P < 0.020$, $\eta^2 = 0.059$) (Fig. 16).

The maximum number of drum beats ($F(2,126) = 14.095$, $P < 0.000$, $\eta^2 = 0.180$; Bonferroni adjusted *post-hoc* tests indicate that the mean difference between K and V is 2.015, $p = 0.0001$; the mean difference between K and D is 1.076, $P = 0.001$) (Fig. 17), and the average number of drum beats ($F(2,126) = 13.442$, $P < 0.000$, $\eta^2 = 0.174$; Bonferroni adjusted *post-hoc* tests indicate that mean difference between K and V is 0.802, $P = 0.0001$; the mean difference between K and D is 0.495, $P = 0.0001$) (Fig. 18) of children and KASPAR per game is significantly lower in condition K than it is in conditions V and D. There is also an interaction

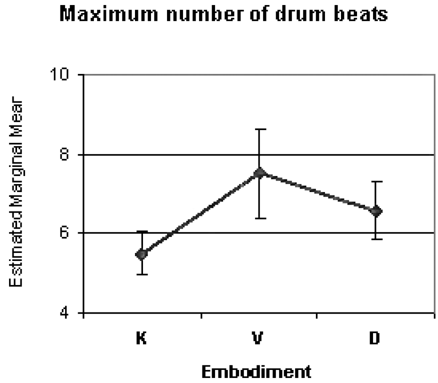

Figure 17. Maximum number of drum beats of children according to the embodiment condition.

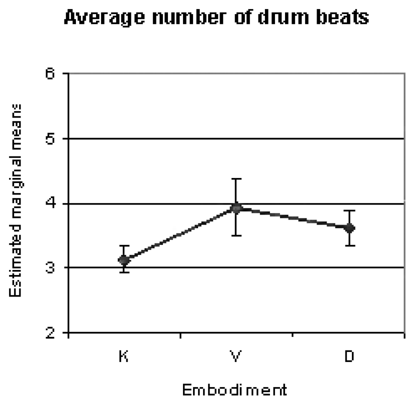

Figure 18. Average number of drum beats of KASPAR according to the embodiment.

effect that suggests that these differences are more pronounced in the presence of gestures, described in Fig. 19 ($F(2,126) = 5.145$, $P < 0.007$, $\eta^2 = 0.074$).

4.3. Effect of the Play Time on the Games and Gestures (H3)

The analysis of the behavioral data showed significant effects for the play time in terms of the drumming and turn-taking performances. In general, it was observed that the participants initially either tried very long and fast patterns or they did not beat the drum loud enough to be detected reliably when they started to play (KASPAR uses a high-level noise filter to filter out high inner noise coming from its joints, so it can only sense loud beats).

The effect of the play time is described in detail in Table 7 (differences between human and KASPAR's perspectives), Table 8 (human's perspective) and Table 9 (KASPAR's perspective). Interestingly, without any external encouragement, as the children played more, it appeared that they got used to the game and were progressively able to synchronize themselves to the robot better. Thus, the error rate (*Errorsum*) decreased significantly over time (Fig. 20) ($F(2,126) = 7.563$, $P < 0.001$,

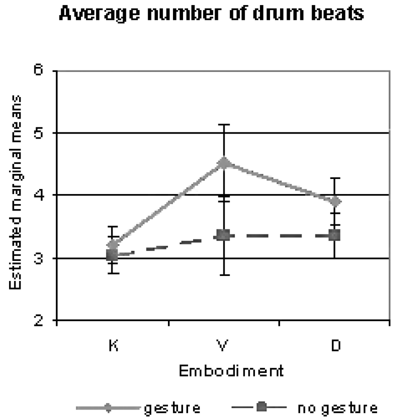

Figure 19. Effect of gestures on the average number of drum beats of humans according to the embodiment condition.

Table 7.
Observed differences between KASPAR-human drumming behaviors according to game order

Order	Difference of sum of beats	Difference of non-zero turns	No. of non-perfect turns	Error in sum of beats	Error in number of turns
1	20.15 ± 16.6	2.53 ± 2.3	15.58 ± 3.8	1.00 ± 0.79	0.79 ± 0.11
2	16.71 ± 13.2	2.64 ± 3.3	15.38 ± 3.7	0.82 ± 0.63	0.77 ± 0.14
3	12.24 ± 10.4	1.82 ± 2.2	14.14 ± 4.6	0.64 ± 0.55	0.74 ± 0.16

Table 8.
Observed human drumming behavior according to game order

Order	Sum of beats	No. of non-zero turns	Maximum no. of beats	Average no. of beats/turn
1	64.15 ± 19.2	19.15 ± 4.57	12	3.44 ± 0.99
2	64.92 ± 16.2	19.56 ± 4.20	13	3.44 ± 1.08
3	66.32 ± 16.8	18.55 ± 4.32	25	3.83 ± 1.71

Table 9.
Observed KASPAR's drumming behavior according to game order

Order	Sum of beats	No. of non-zero turns	Maximum no. of beats	Average no. of beats/turn
1	47.52 ± 15.9	17.86 ± 3.53	14	2.70 ± 0.98
2	52.33 ± 16.5	18.02 ± 3.77	13	2.96 ± 1.07
3	60.05 ± 15.9	18.03 ± 4.06	23	3.56 ± 1.66

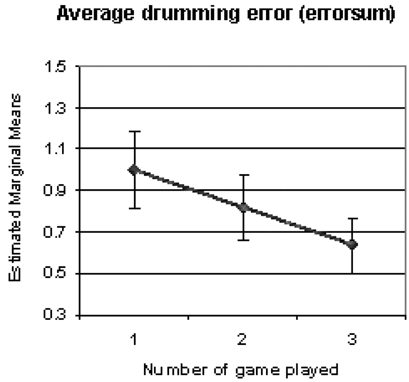

Figure 20. Average errors in sum of beats according to order.

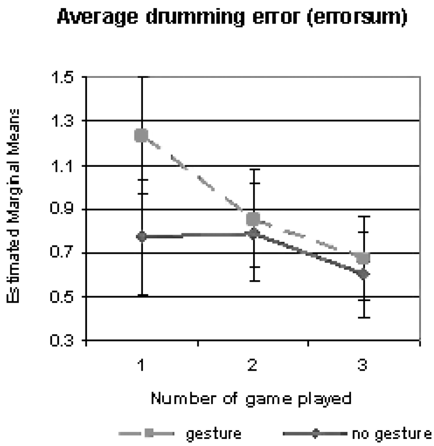

Figure 21. Effect of gestures on the average errors in sum of beats according to order.

$\eta^2 = 0.106$; Bonferroni adjusted *post-hoc* tests indicate that the mean difference between the first game and the third game is 0.365, $P = 0.003$), implying a significant interaction effect between order and gestures, suggesting that this effect is more pronounced in the gesture condition. ($F(2,126) = 2.952$, $P < 0.056$, $\eta^2 = 0.044$), as shown in Fig. 21. Similarly, the difference between KASPAR's and human's total drumming decreases as the children play more games ($F(2,126) = 7.067$; $P < 0.001$, $\eta^2 = 0.099$; Bonferroni adjusted *post-hoc* tests indicate that the mean difference between the first game and the third game is 7.242, $P = 0.005$; the mean difference between the second and the third game is 5.106, $P = 0.008$). There is also an significant decrease of the error rate in the turn-taking (*Errorturn*) between the first and third games ($F(2,126) = 5.520$, $P < 0.022$, $\eta^2 = 0.079$) (Fig. 22) that supports hypothesis H3.

As shown in the Fig. 23, the maximum number of beats per game increased significantly between the first and third game ($F(2,126) = 7.455$, $P < 0.001$, $\eta^2 =$

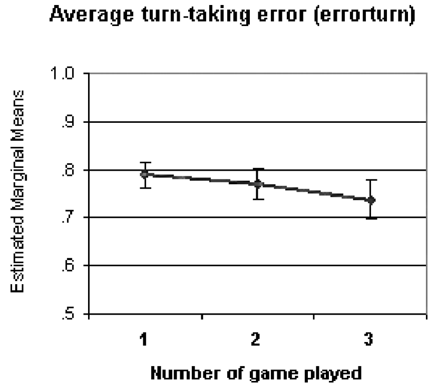

Figure 22. Average errors in number of turns according to game order.

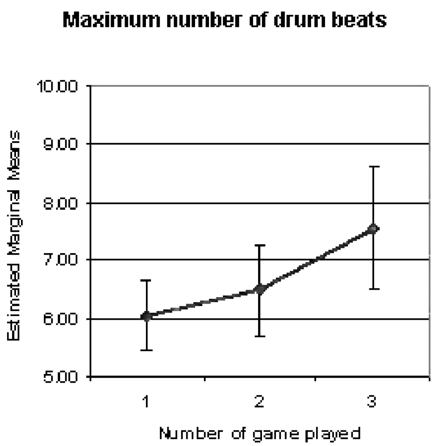

Figure 23. Maximum number of drum beats of children per turn according to order.

0.079; Bonferroni adjusted *post-hoc* tests indicate that the mean difference between the first game and the third game is 1.33, $P = 0.013$). Also, the average drumming per turn increases significantly in the third game ($F(2,126) = 4.732$, $P < 0.010$, $\eta^2 = 0.069$; Bonferroni adjusted *post-hoc* tests indicate that the mean difference between the first and the third game approaches significance, the mean difference is 0.428, $P = 0.06$) (Fig. 24), which may suggest that participants played more beats, possibly due to a stronger involvement in the game, as they played more games.

The number of non-zero turns differs significantly for the children according to order as shown in Fig. 25 ($F(2,126) = 3.800$, $P < 0.025$, $\eta^2 = 0.056$; Bonferroni adjusted *post-hoc* tests indicate that the mean difference between the second and the third game is 1.182, $P = 0.011$). Children (and consequently KASPAR) played in less turns with a higher number of beats per turn and with longer durations when the second and last games were compared.

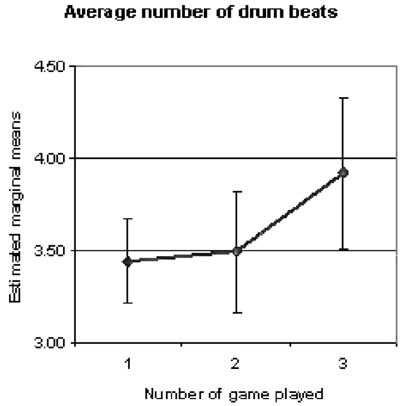

Figure 24. Average number of drum beats of children according to order.

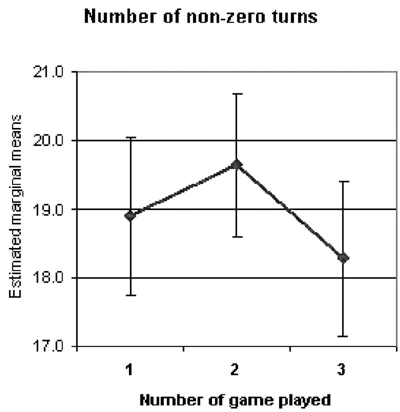

Figure 25. Number of the non-zero turns of children per game according to order.

4.4. Effect of Gender

A significant effect was found for gender in the evaluation of the robot's social attraction. This effect is described in Fig. 26 and suggests that male children found the robot more socially attractive in the physical embodiment condition than in the two other embodiment conditions, while a similar effect is not evident for females ($F(2,126) = 3,06$, $P < 0.051$, $\eta^2 = 0.046$) (Fig. 26).

In terms of behavioral data, gender showed significant differences. When different embodiment conditions were compared, there was a significant interaction effect between embodiment and gender in terms of the maximum number of beats played per turn ($F(2,126) = 8.497$, $P < 0.000$, $\eta^2 = 0.117$). The effect, shown in Fig. 27, suggest that when the children play with the physical robot, their performance is similar, but in the absence of the physical robot male children play more beats than the female children. This effect is most pronounced in the virtual embodiment condition. There could possibly be a link between the males tending to play computer games and these results. They may view the game with the two-dimensional pro-

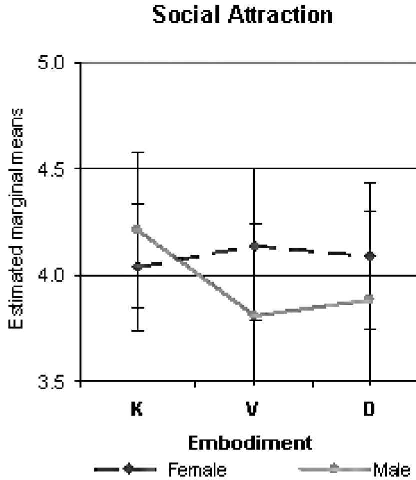

Figure 26. Gender differences in the robot's social attraction in the different embodiment conditions.

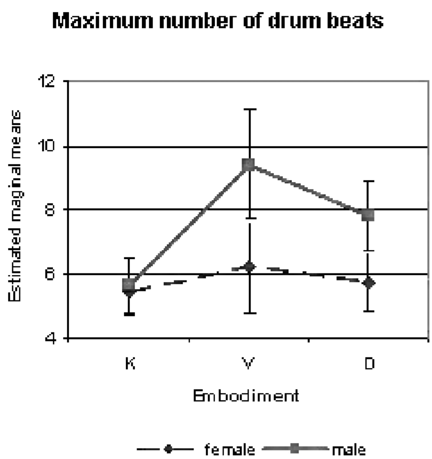

Figure 27. Effect of gender on the maximum number of drum beats of children according to the embodiment condition.

jected image of the robot (V condition) like a computer game and this might affect their performances. Further analysis is needed to investigate this issue.

In terms of the play time effect, the gender differences are significant in the number of drum beats played. There is a significant main effect for gender ($F(1,63) = 7.042$, $P = 0.01$, $\eta^2 = 0.099$; due to there only being two levels for gender, a Bonferroni test was not conducted), for game order ($F(2,126) = 7.455$, $P = 0.001$, $\eta^2 = 0.104$; Bonferroni adjusted *post-hoc* test found the following significant mean differences: first and third game, mean difference $= 1.493$, $P = 0.004$; second and third, mean difference 1.060, $P = 0.028$), as well as a significant interaction between gender and game order ($F(2,126) = 4.639$, $P < 0.011$, $\eta^2 = 0.068$). This interaction effect is described in Fig. 28 and suggests that the male participants

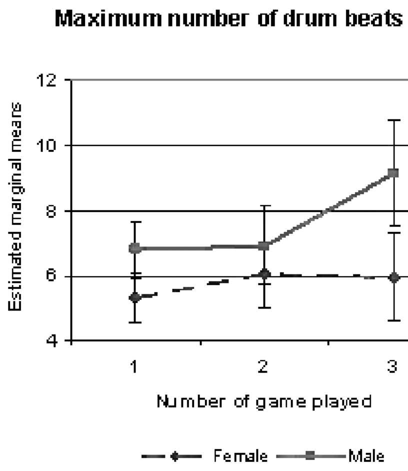

Figure 28. Effect of gender on the maximum number of drum beats of children according to the sequential order.

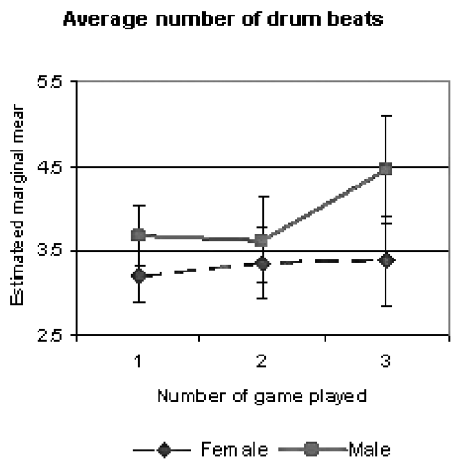

Figure 29. Effect of gender on the average number of drum beats of humans according to the sequential order.

have a more pronounced difference between the third game and the other games than the females.

For average number of beats, a significant main effect was found for gender ($F(1,63) = 5.212$, $P = 0.026$, $\eta^2 = 0.075$; Bonferroni test not conducted due to gender only having two levels) and game order ($F(2,126) = 4.732$, $P = 0.012$, $\eta^2 = 0.069$; Bonferroni adjusted *post-hoc* tests found a significant mean difference between the first and the the third game, mean difference $= 0.481$, $P = 0.035$); there was also an interaction effect between gender and game order approaching significance ($F(2,126) = 2.922$, $P < 0.057$, $\eta^2 = 0.044$). This interaction effect is described in Fig. 29, and suggests that the difference between the first and the third game is more pronounced for the male participants.

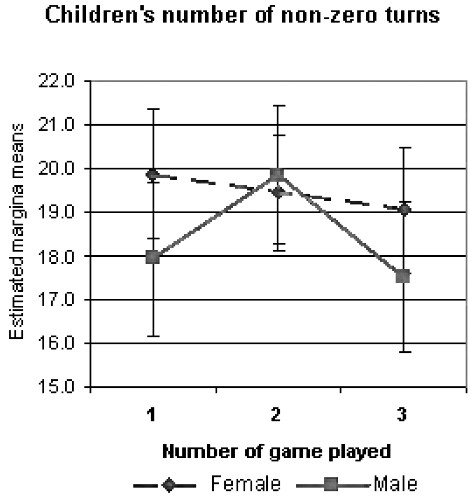

Figure 30. Effect of gender on the number of non-zero turns of humans according to game order.

For the number of turns played, a significant main effect was found for game order ($F(2,126) = 3.800$, $P = 0.025$, $\eta^2 = 0.056$; Bonferroni adjust *post-hoc* test found a significant mean difference between the the second and the third game, mean difference $= 1.359$, $P = 0.002$). However, an interaction effect was found for gender and game order ($F(2,126) = 3.263$, $P < 0.041$, $\eta^2 = 0.049$); this effect is described in Fig. 30 and suggests that this difference is primarily due to the behavior of the male participants.

5. Conclusions

In this research, we studied the effect of embodiment and gestures on a human–humanoid drumming game. We tested different levels of embodiment of the humanoid robot which autonomously played games with child participants. Half of the children interacted with a robot that made simple head gestures while imitating the child's drumming, while the rest of them played with a robot that did not make any gestures, but simply played its drum.

The analysis of results from video recordings, questionnaire data and the robot's recordings of the behavioral data gave either partial or full support for our original hypotheses as formulated in Section 2. The physical embodied robot (K) in the gesture condition has been evaluated by the children as the interaction that they enjoyed the most. The drumming performance of the child–robot pair is the highest in the physical robot condition, and decreases in the virtual and disembodied robot conditions. Similarly, best turn-taking was achieved when they played with the physical robot; their coordination got worse in the virtual and disembodied robot conditions.

Results of questionnaire data analysis and behavioral data support the expectation that embodiment can play an important role in social interaction tasks. In particular, our child participants found the presence of a physical robot most enjoyable and believed it to be more pleasant to play with than a hidden robot or a virtual robot.

Overall, the questionnaire results show that children's opinions are not effected by the embodiment conditions, nor by the presence or absence of robot gestures. Nevertheless, it is interesting to note that the data indicates a trend in which the children generally appreciated more the interaction with the physically embodied robot than the other two conditions. Despite of that small result, significant interaction effects of embodiment and gesture conditions have been found in terms of enjoyment, intelligence and appearance.

In terms of enjoyment, there is a significant difference between the embodiment conditions when the robot made gestures during its drumming, where participants in the gesture condition enjoyed interacting with the embodied robot more than with the two other embodiment conditions. The result concerning intelligence (in which the perceived intelligence of the robot was less for the video and disembodied condition than for the embodied condition if gestures were used, while the opposite was true when no-gestures were performed) is also quite interesting, possibly highlighting the importance of physical embodiment for effective use of and the processing of non-verbal cues in social interactions. The result regarding appearance might suggest that the gestures used by the robot might be appropriate for the projected image, but not smooth enough for a robot sat in front of a child.

Moreover, the behavioral data of the children and the robot support that there is a significant difference between the embodiment conditions as we hypothesized. The presence of the physically embodied robot motivated the children positively, and helped to improve the turn-taking and drumming between the robot and the children significantly. When the robot made gestures whilst drumming, the differences between the drumming and turn-taking performances belonging to different embodiment conditions increased significantly. Gestures played a positive role, especially when the child played with the physically embodied robot in terms of turn-taking.

Also, there is a significant difference in the error rates of drumming and turn-taking between the first and the third games. Thus, as hypothesized, the children enjoyed the game more, and both the drumming and the turn-taking performance of the robot and the children improved as they played more. When the gestures were introduced, especially in the first game, there was a significant difference in the turn-taking errors — as the children played more, they got used to the robot and the gestures, and this difference decreased.

Note that we are aware of the limitations of our study. For example, the variances within the sample were quite large in comparison to the effect sizes. This is to be expected — the novelty of HRI scenarios, and considering that the sample consisted of children, would lead one to expect that idiosyncracies of the individual participant would impact both interactions and evaluations of these. Owing to this, a sample size such as the one presented in the paper (which is quite large compared to similar studies) is a reasonable and accepted way of controlling for such idiosyncracies in order to avoid a type II error ('Failing to reject a false null hypothesis') [47].

Several directions for future work can be envisaged. Considering that play is often a 'group activity', one may study teams of children playing together with the robot in order to investigate how they interact with each other as well as with the robot. Such 'social facilitation effects' in HRI have been found, for example, in our previous studies involving children playing games with a non-humanoid mobile robot [48]. Also, the experience of the human–humanoid drumming game could be compared with a human–human drumming game. Furthermore, adding visual feedback with the use of the robot's internal cameras may be helpful in enhancing the robot's interaction with the children. In that way the robot could adapt itself to the behavioral changes in the children to achieve a higher level of social interaction with them. Finally, different robot appearances (e.g., humanoid *versus* mechanical-looking or zoomorphic) and different robot behaviors could be varied systematically in studies comparing virtual and physical embodiment conditions.

On a general note, recently more emphasis has been given to the use of virtual social agents (e.g., virtual characters) in the context of interaction with humans [49–52]. Graphical characters in virtual environments (*cf*; computer games) can create scenarios far more complex than typical HRI scenarios, which explains their popularity in entertainment and educational applications, for example. However, interaction with virtual characters typically requires the use of dedicated interfaces, e.g., keyboard, mouse, Wii, etc., while interactions with robots, as described in this paper, do not necessarily require specialized input devices. As the results in this paper indicate, there is a clear benefit of using physical embodied characters, compared to virtual characters, as far as children's responses are concerned. Thus, despite limitations concerning the complexity of autonomous intelligent behavior and interactive capabilities of state-of-the-art robots, our research supports the need for physically embodied interaction in suitable scenarios. Note that in therapy applications the physical dimension of interaction can provide an additional strong incentive and therapeutic objective, and not unsurprisingly, more and more embodied robots have been used in special education, engaging children with special needs in meaningful interactions [53–55].

This large-scale study with an autonomous humanoid robot and children is one of the first studies in this domain of comparing physical and virtual robot embodiments. The achievements and findings here suggest implications for research in application areas involving robots and children. Our results indicate that the embodiment of the robot (virtual or physical) has a significant impact on the objective performance and children's subjective evaluation of the interaction.

Acknowledgements

This work was conducted within the EU Integrated Project RobotCub (Robotic Open-architecture Technology for Cognition, Understanding and Behaviours) funded by the EC through the E5 Unit (Cognition) of FP6-IST under Contract FP6-004370. We would like to thank our colleagues in the Adaptive Systems Research Group for their help in organizing the experiments: Scott Watson, Josh Wainer, Qiming Shen, Dr Michael L. Walters and Dr Joe Saunders. We would like to thank Dr René te Boekhorst for suggestions concerning the data analysis and Dr Frank Broz for proofreading the manuscript. Also, we are thankful to the schools who attended our event and the teachers who voluntarily helped us during the experiments.

References

1. K. Shinozaki, A. Iwatani and R. Nakatsu, Construction and evaluation of a robot dance system, in: *Proc. 12th Int. Conf. on Knowledge-based Intelligent Information and Engineering Systems*, Zagreb, pp. 682–691 (2008).
2. A. Tiago, C. Martinho, I. Leite and A. Paiva, iCat, the chess player: the influence of embodiment in the enjoyment of a game, in: *Proc. 7th Int. Conf. on Autonomous Agents and Multiagent Systems*, Estoril, pp. 1253–1256 (2008).
3. D. Golubovic, B. Li and H. Hu, A hybrid software platform for Sony AIBO robots, in: *Proc. RoboCup 2003: Robot Soccer World Cup VII*, Padua, pp. 478–486 (2003).
4. F. Michaud and S. Caron, Roball — an autonomous toy-rolling robot, in: *Proc. Workshop on Interactive Robot Entertainment*, Pittsburgh, PA, pp. 127–134 (2000).
5. A. Tapus and M. J. Matarić, Towards socially assistive robotics, *Int. J. Robotics Soc. Japan* **24**, 576–578 (2006).
6. K. Inoue, K. Wada and Y. Ito, Effective application of paro: seal type robots for disabled people according to ideas of occupational therapists, in: *Proc. ICCHP*, Linz, pp. 1321–1324 (2008).
7. B. Robins, K. Dautenhahn and J. Dubowsky, Does appearance matter in the interaction of children with autism with a humanoid robot? *Interact. Studies* **7**, 479–512 (2006).
8. K. Wada and T. Shibata, Robot therapy in a care house — its sociopsychological and physiological effects on the residents, in: *Proc. Int. Conf. on Robotics and Automation*, Orlando, FL, pp. 3966–3971 (2006).
9. I. Werry, K. Dautenhahn, B. Ogden and W. Harwin, Can social interaction skills be taught by a social agent? The role of a robotic mediator in autism therapy, in: *Proc. Int. Conf. on Cognitive Technology*, Coventry, pp. 57–74 (2001).
10. T. W. Bickmore and J. Cassell, Social dialogue with embodied conversational agents, in: *Natural, Intelligent and Effective Interaction with Multimodal Dialogue Systems*, J. v. Kuppevelt, L. Dybkjaer and N. Bernsen (Eds), pp. 23–54. Kluwer Academic, New York, NY (2005).
11. T. W. Bickmore and R. W. Picard, Establishing and maintaining long-term human computer relationships, *ACM Trans. Comp.-Hum. Interact.* **12**, 293–327 (2005).
12. K. M. Lee, N. Park and H. Song, Can a robot be perceived as a developing creature?: effects of artificial developments on social presence and social responses toward robots in human–robot interaction, Paper presented at: *Int. Communication Association Conf.*, New Orleans, LA (2004).
13. T. Fong, I. Nourbakhsh and K. Dautenhahn, A survey of socially interactive robots, *Robotics Autonomous Syst.* **42**, 143–166 (2003).
14. C. Lathan, J. Vice, M. Tracey, C. Plaisant, A. Druin, K. Edward and J. Montemayor, Therapeutic play with a storytelling robot, in: *Proc. Conf. on Human Factors in Computing Systems*, Seattle, WA, pp. 27–28 (2001).

15. J. Wainer, D. J. Feil-Seifer, D. A. Shell and M. J. Mataric, Embodiment and human–robot interaction: a taskbased perspective, in: *Proc. Int. Conf. on Human–Robot Interaction*, Jeju Island, pp. 872–877 (2007).
16. T. Quick, K. Dautenhahn, C. Nehaniv and G. Roberts, Understanding embodiment, system–environment coupling and the emergence of adaptive behaviour, in: *Intelligence Artificielle Située*, A. Drogoul and J.-A. Meyer (Eds), pp. 13–31. Hermes Science, Paris (1999).
17. K. Dautenhahn, B. Ogden and T. Quick, From embodied to socially embedded agents — implications for interaction-aware robots, *Cognit. Syst. Res.* **3**, 397–428 (2002).
18. R. L. Chrisley and T. Ziemke, Embodiment, in: *Encyclopedia of Cognitive Science*, L. Nadel (Ed.), pp. 1102–1108. Macmillan, London (2002).
19. K. M. Lee, Y. Jung, J. Kim and S. R. Kim, Are physically embodied social agents better than, disembodied social agents?: the effects of physical embodiment, tactile interaction, and people's loneliness in human–robot interaction, *Int. J. Hum.-Comput. Stud.* **64**, 962–973 (2006).
20. C. Bartneck, eMuu — an embodied emotional character for the ambient intelligent home, *PhD Thesis*, Eindhoven (2002).
21. A. Tapus and M. J. Matarić, Socially assistive robotic music therapist for maintaining attention of older adults with cognitive impairments, in: *Proc. AAAI Fall Symp. AI in Eldercare: New Solutions to Old Problem*, Washington, DC, pp. 297–298 (2008).
22. J. Wainer, D. J. Feil-Seifer, D. A. Shell and M. J. Mataric, The role of physical embodiment in human–robot interaction, in: *Proc. IEEE Int. Workshop on Robot and Human Interactive Communication*, Hatfield, pp. 117–122 (2006).
23. A. Powers, S. Kiesler, S. Fussell and C. Torrey, Comparing a computer agent with a humanoid robot, in: *Proc. ACM/IEEE Int. Conf. on Human–Robot Interaction*, Washington, DC, pp. 145–152 (2007).
24. C. Breazeal, *Designing Sociable Robots*. MIT Press, Cambridge, MA (2002).
25. C. Breazeal, Toward sociable robots, *Robotics Autonomous Syst.* **42**, 167–175 (2003).
26. C. Crick, M. Munz and B. Scassellati, Synchronization in social tasks: robotic drumming, in: *Proc. IEEE Int. Symp. on Robot and Human Interactive Communication*, Hatfield, pp. 97–102 (2006).
27. G. Weinberg and S. Driscoll, Robot–human interaction with an anthropomorphic percussionist, in: *Proc. Int. ACM Computer Human Interaction Conf.*, Montreal, pp. 1229–1232 (2006).
28. S. Kotosaka and S. Schaal, Synchronized robot drumming by neural oscillator, in: *Proc. Int. Symp. on Adaptive Motion of Animals and Machines*, Montreal, QC (2000).
29. H. Kose-Bagci, K. Dautenhahn, C. L. Nehaniv, Emergent dynamics of turn-taking interaction in drumming games with a humanoid robot, in: *Proc. IEEE Int. Symp. on Robot and Human Interactive Communication*, Munich, pp. 346–353 (2008).
30. H. Kose-Bagci, K. Dautenhahn, D. S. Syrdal and C. L. Nehaniv, Drum-mate: A Human-Humanoid Drumming Experience, in: *Proc. IEEE-RAS Int. Conf. on Humanoid Robots*, Pittsburgh, PA, pp. 242–247 (2007).
31. M. P. Blow, K. Dautenhahn, A. Appleby, C. Nehaniv and D. Lee, Perception of robot smiles and dimensions for human–robot interaction design, in: *Proc. IEEE Int. Symp. on Robot and Human Interactive Communication*, Hatfield, pp. 469–474 (2006).
32. P. Schermerhorn, M. Scheutz and C. R. Crowell, Robot social presence and gender: do females view robots differently than males?, in: *Proc. ACM/IEEE Int. Conf. on Human–Robot Interaction*, Amsterdam, pp. 263–270 (2008).
33. K. Dautenhahn, M. Walters, S. Woods, K. L. Koay, C. L. Nehaniv, A. Sisbot, R. Alami and T. Siméon, How may I serve you?: a robot companion approaching a seated person in a help-

ing context, in: *Proc. ACM SIGCHI/SIGART Conf. on Human–Robot Interaction*, Salt Lake City, UT, pp. 172–179 (2006).

34. D. S. Syrdal, K.-L. Koay, M. L. Walters and K. Dautenhahn, A personalised robot companion? The role of individual differences on spatial preferences in HRI scenarios, in: *Proc. IEEE Int. Symp. on Robot and Human Interactive Communication*, Jeju Island, pp. 26–29 (2007).
35. P. Schermerhorn, M. Scheutz and C. R. Crowell, Robot social presence and gender: do females view robots differently than males?, in: *Proc. ACM/IEEE Int. Conf. on Human–Robot Interaction*, Amsterdam, pp. 263–270 (2008).
36. FearNot! event, available from: http://www.herts.ac.uk/news-and-events/latest-news/Antibullying-software-and-robots.cfm, last accessed: 3 June 2009.
37. S. Sabanovic, M. P. Michalowski and R. Simmons, Robots in the wild: observing human–robot social interaction outside the lab, in: *Proc. AMC 2006*, Antalya, pp. 576–581 (2006).
38. T. Kanda, M. Shiomi, Z. Miyashita, H. Ishiguro and N. Hagita, An affective guide robot in a shopping mall, in: *Proc. ACM/IEEE Int. Conf. on Human–Robot Interaction*, San Diego, CA, pp. 173–180 (2009).
39. I. R. Nourbakhsh, C. Kunz and T. Willeke, The mobot museum robot installations: a five year experiment, in: *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Las Vegas, NV, pp. 3636–3641 (2003).
40. N. Dahlback, A. Jonsson and L. Ahrenberg, Wizard of Oz studies — why and how, in: *Proc. Int. Conf. on Intelligent User Interfaces*, Orlando, FL, pp. 193–200 (1993).
41. B. Robins, K. Dautenhahn and P. Dickerson, From isolation to communication: a case study evaluation of robot assisted play for children with autism with a minimally expressive humanoid robot, in: *Proc. Int. Conf. on Advances in Computer–Human Interactions*, Cancun, in press (2009).
42. B. Robins, K. Dautenhahn, R. te Boekhorst and C. L. Nehaniv, Behaviour delay and robot expressiveness in child–robot interactions: a user study on interaction kinesics, in: *Proc. ACM/IEEE Int. Conf. on Human–Robot Interaction*, Amsterdam, pp. 17–24 (2008).
43. G. Metta, P. Fitzpatrick and L. Natale, YARP: yet another robot platform, *Int. J. Adv. Robotics Syst. (Special Issue on Software Development and Integration in Robotics)* **3**, 43–48 (2006).
44. R. Bencina and P. Burk, Portaudio, available from: http://www.portaudio.com/trac/wiki/
45. H. Kose and H. L. Akin, Object recognition in robot football using a one dimensional image, in: *Proc. Turkish Symp. on Artificial Intelligence and Neural Networks*, Antalya, pp. 291–300 (2001).
46. J. C. McCroskey and T. A. McCain, The measurement of interpersonal attraction, *Speech Monogr.* **41**, 261–266 (1974).
47. J. Cohen, A power primer, *Psychol. Bull.* **112**, 155–159 (1992).
48. S. Woods, K. Dautenhahn and C. Kaouri, Is someone watching me? Consideration of social facilitation effects in human–robot interaction experiments, in: *Proc. IEEE Int. Symp. on Computational Intelligence in Robotics and Automation*, Espoo, pp. 53–60 (2005).
49. J. Cassell, T. Bickmore, H. Vilhjálmsson and H. Yan, More than just a pretty face: affordances of embodiment, in: *Proc. Int. Conf. on Intelligent User Interfaces*, New Orleans, LA, pp. 52–59 (2000).
50. J. Rickel, Intelligent virtual agents for education and training: opportunities and challenges, *Lecture Notes Comput. Sci.* **2190**, 15–22 (2001).
51. L. Chiattaro, I. Ieronutti and S. Rigutti, Supporting presentation techniques based on virtual humans in educational virtual worlds, in: *Proc. Int. Conf. on Cyberworlds*, Singapore, pp. 245–252 (2005).

52. M. Sapouna, D. Wolke, N. Vannini, S. Watson, S. Woods, W. Schneider, S. Enz, L. Hall, A. Paiva, E. Andre, K. Dautenhahn and R. Aylett, Virtual learning intervention to reduce bullying victimization in primary school: a controlled trial, *J. Child Psychol. Psychiatr.*, in press (2009).
53. A. Cook, K. Howery, J. Gu and M. Meng, Robot enhanced interaction and learning for children with profound physical disabilities, *Technol. Disabil.* **13**, 1–8 (2000).
54. B. Robins, P. Dickerson and K Dautenhahn, Robots as embodied beings — interactionally sensitive body movements in interactions among autistic children and a robot, in: *Proc. IEEE Int. Workshop on Robot and Human Interactive Communication*, Nashville, TN, pp. 54–59 (2005).
55. F. Michaud, T. Salter, A. Duquette, H. Mercier, H. Larouche and F. Larose, Assistive technologies and child–robot interaction, in: *Proc. AAAI Spring Symp. on Multidisciplinary Collaboration for Socially Assistive Robotics*, Stanford, CA (2007).

Appendix A

Descriptive Data Related to Questionnaire Data

Table A.1.
Descriptive data of the embodiment effect

	Embodiment condition	Mean	SE	95% Confidence interval	
				Lower bound	Upper bound
Enjoyment	*K*	4.457	0.090	4.278	4.637
	V	4.262	0.125	4.013	4.511
	D	4.376	0.097	4.182	4.570
Performance	*K*	3.667	0.139	3.389	3.944
	V	3.735	0.158	3.418	4.051
	D	3.735	0.142	3.451	4.019
Appearance	*K*	4.212	0.099	4.014	4.410
	V	4.020	0.118	3.784	4.256
	D	3.985	0.105	3.775	4.195
Social attraction	*K*	4.111	0.113	3.884	4.338
	V	4.009	0.139	3.730	4.287
	D	4.008	0.133	3.742	4.273
Intelligence	*K*	4.325	0.105	4.114	4.535
	V	4.275	0.115	4.045	4.505
	D	4.274	0.123	4.028	4.520
Involvement	*K*	4.455	0.087	4.280	4.629
	V	4.340	0.104	4.132	4.547
	D	4.330	0.107	4.117	4.543

Table A.2.
Descriptive data of the gesture effect

	Gesture condition	Mean	SE	95% Confidence interval	
				Lower bound	Upper bound
Enjoyment	gesture	4.422	0.120	4.181	4.662
	no-gesture	4.308	0.118	4.071	4.545
Performance	gesture	3.581	0.164	3.253	3.908
	no-gesture	3.843	0.164	3.516	4.171
Appearance	gesture	4.040	0.118	3.804	4.276
	no-gesture	4.104	0.118	3.868	4.340
Social attraction	gesture	4.135	0.162	3.811	4.460
	no-gesture	3.949	0.160	3.630	4.269
Intelligence	gesture	4.219	0.123	3.973	4.464
	no-gesture	4.364	0.121	4.122	4.605
Involvement	gesture	4.401	0.120	4.161	4.641
	no-gesture	4.348	0.118	4.112	4.585

Table A.3.
Descriptive data of the gender effect

	Gender	Mean	SE	95% Confidence interval	
				Lower bound	Upper bound
Enjoyment	male	4.359	0.134	4.091	4.627
	female	4.368	0.109	4.149	4.586
Performance	male	3.519	0.180	3.158	3.879
	female	3.846	0.150	3.546	4.146
Appearance	male	4.000	0.130	3.740	4.260
	female	4.123	0.108	3.906	4.339
Social attraction	male	3.968	0.181	3.607	4.329
	female	4.090	0.147	3.795	4.384
Intelligence	male	4.314	0.133	4.048	4.580
	female	4.415	0.109	4.197	4.632
Involvement	male	4 269	0.137	3995	4 543
	female	4.308	0.112	4.084	4.531

Table A.4.
Descriptive data of the interaction between embodiment and gesture

	Embodiment condition	Mean	SE	95% Confidence interval	
				Lower bound	Upper bound
Enjoyment[a]					
gesture	*K*	4.688	0.128	4.432	4.943
	V	4.281	0.178	3.927	4.636
	D	4.297	0.138	4.020	4.573
no-gesture	*K*	4.227	0.126	3.976	4.479
	V	4.242	0.175	3.893	4.592
	D	4.455	0.136	4.182	4.727
Performance					
gesture	*K*	3.621	0.196	3.229	4.014
	V	3.682	0.224	3.234	4.129
	D	3.439	0.201	3.038	3.841
no-gesture	*K*	3.712	0.196	3.320	4.105
	V	3.788	0.224	3.340	4.235
	D	4.030	0.201	3.629	4.432
Appearance[a]					
gesture	*K*	4.162	0.140	3.881	4.442
	V	4.182	0.167	3.848	4.515
	D	3.778	0.149	3.480	4.075
no-gesture	*K*	4.263	0.140	3.982	4.543
	V	3.859	0.167	3.525	4.192
	D	4.192	0.149	3.894	4.489
Social attraction					
gesture	*K*	4.328	0.162	4.005	4.651
	V	4.078	0.198	3.681	4.475
	D	4.000	0.189	3.622	4.378
no-gesture	*K*	3.894	0.159	3.576	4.212
	V	3.939	0.195	3.549	4.330
	D	4.015	0.186	3.643	4.388
Involvement					
gesture	*K*	4.500	0.124	4.251	4.749
	V	4.422	0.148	4.126	4.717
	D	4.281	0.152	3.978	4.585
no-gesture	*K*	4.409	0.123	4.164	4.654
	V	4.258	0.146	3.966	4.549
	D	4.379	0.150	4.080	4.678
Intelligence[a]					
gesture	*K*	4.438	0.150	4.138	4.737
	V	4.156	0.164	3.829	4.484
	D	4.063	0.176	3.712	4.413
no-gesture	*K*	4.212	0.148	3.917	4.507
	V	4.394	0.161	4.072	4.716
	D	4.485	0.173	4.139	4.830

[a] Significant interaction effect.

Table A.5.
Descriptive data of the interaction between embodiment and gender

	Embodiment condition	Mean	SE	95% Confidence interval	
				Lower bound	Upper bound
Enjoyment					
male	*K*	4.442	0.149	4.145	4.740
	V	4.269	0.197	3.876	4.663
	D	4.365	0.154	4.057	4.674
female	*K*	4.462	0.122	4.218	4.705
	V	4.256	0.161	3.935	4.578
	D	4.385	0.126	4.133	4.636
Performance					
male	*K*	3.352	0.211	2.930	3.774
	V	3.556	0.246	3.064	4.047
	D	3.648	0.229	3.190	4.106
female	*K*	3.885	0.176	3.533	4.236
	V	3.859	0.205	3.450	4.268
	D	3.795	0.191	3.414	4.176
Appearance					
male	*K*	4.198	0.155	3.887	4.508
	V	3.889	0.186	3.517	4.261
	D	3.914	0.169	3.576	4.251
female	*K*	4.222	0.129	3.964	4.480
	V	4.111	0.155	3.802	4.420
	D	4.034	0.141	3.753	4.315
Social attraction[a]					
male	*K*	4.212	0.184	3.844	4.579
	V	3.808	0.218	3.372	4.244
	D	3.885	0.209	3.467	4.302
female	*K*	4.038	0.150	3.739	4.338
	V	4.141	0.178	3.785	4.497
	D	4.090	0.171	3.749	4.431
Involvement					
male	*K*	4.462	0.138	4.185	4.738
	V	4.250	0.164	3.922	4.578
	D	4.231	0.168	3.895	4.566
female	*K*	4.449	0.113	4.223	4.674
	V	4.397	0.134	4.129	4.666
	D	4.397	0.137	4.123	4.671
Intelligence					
male	*K*	4.462	0.166	4.129	4.794
	V	4.231	0.183	3.865	4.597
	D	4.115	0.197	3.721	4.510
female	*K*	4.231	0.136	3.959	4.502
	V	4.308	0.150	4.009	4.606
	D	4.385	0.161	4.062	4.707

[a] Significant interaction effect.

Descriptive Data Related to Behavioral Data

Table A.6.
Descriptive data of the embodiment effect

	Embodiment condition	Mean	SE	95% Confidence interval	
				Lower bound	Upper bound
Errorsum[a]	*K*	0.623	0.063	0.497	0.750
	V	0.938	0.082	0.774	1.102
	D	0.899	0.095	0.710	1.088
Errorturn	*K*	0.741	0.016	0.709	0.772
	V	0.778	0.017	0.744	0.812
	D	0.780	0.017	0.746	0.814
Diffsum[a]	*K*	12.530	1.356	9.822	15.239
	V	18.379	1.622	15.139	21.619
	D	18.197	2.011	14.180	22.214
Maxofbeats[a]	*K*	5.500	0.261	4.978	6.022
	V	7.515	0.571	6.375	8.655
	D	6.576	0.367	5.843	7.308
Avgofbeats[a]	*K*	3.129	0.101	2.927	3.332
	V	3.932	0.219	3.494	4.369
	D	3.624	0.130	3.365	3.883

[a] Significant interaction effect.

Table A.7.
Descriptive data of the interaction between embodiment and gender

	Embodiment condition	Mean	SE	95% Confidence interval	
				Lower bound	Upper bound
Maxofbeats[a]					
male	*K*	5.630	0.408	4.814	6.445
	V	9.407	0.841	7.727	11.088
	D	7.815	0.542	6.732	8.898
female	*K*	5.410	0.340	4.732	6.089
	V	6.205	0.700	4.807	7.603
	D	5.718	0.451	4.817	6.619

[a] Significant interaction effect.

Table A.8.
Descriptive data of the interaction between embodiment and gesture

	Embodiment condition	Mean	SE	95% Confidence interval	
				Lower bound	Upper bound
Errorsum[a]					
gesture	*K*	0.642	0.090	0.463	0.822
	V	0.977	0.116	0.745	1.208
	D	1.139	0.134	0.872	1.406
no-gesture	*K*	0.604	0.090	0.425	0.784
	V	0.899	0.116	0.667	1.130
	D	0.659	0.134	0.392	0.926
Errorturn[a]					
gesture	*K*	0.705	0.023	0.659	0.750
	V	0.820	0.024	0.772	0.868
	D	0.820	0.024	0.773	0.868
no-gesture	*K*	0.777	0.023	0.732	0.822
	V	0.735	0.024	0.687	0.783
	D	0.740	0.024	0.692	0.788
Diffsum[a]					
gesture	*K*	11.636	1.917	7.806	15.467
	V	16.606	2.294	12.024	21.188
	D	21.485	2.843	15.805	27.165
no-gesture	*K*	13.424	1.917	9.594	17.254
	V	20.152	2.294	15.569	24.734
	D	14.909	2.843	9.229	20.589
Avgofbeats[a]					
gesture	*K*	3.211	0.143	2.925	3.497
	V	4.510	0.310	3.891	5.129
	D	3.897	0.183	3.531	4.264
no-gesture	*K*	3.048	0.143	2.761	3.334
	V	3.353	0.310	2.734	3.972
	D	3.351	0.183	2.985	3.717

[a] Significant interaction effect.

Table A.9.
Descriptive data of the game order effect

	Game order	Mean	SE	95% Confidence interval	
				Lower bound	Upper bound
Errorsum[a]	1	1.002	0.094	0.815	1.190
	2	0.821	0.079	0.663	0.978
	3	0.638	0.068	0.502	0.773
Errorturn[a]	1	0.790	0.014	0.761	0.818
	2	0.770	0.017	0.736	0.803
	3	0.738	0.020	0.698	0.778
Diffsum[a]	1	19.940	2.095	15.755	24.125
	2	17.745	1.774	14.201	21.289
	3	12.826	1.338	10.153	15.499
Maxofbeats[a]	1	6.056	0.295	5.466	6.645
	2	6.489	0.395	5.700	7.277
	3	7.548	0.525	6.499	8.598
Avgofbeats[a]	1	3.439	0.116	3.208	3.670
	2	3.491	0.163	3.165	3.817
	3	3.920	0.206	3.509	4.331
Non-zeroturns[a]	1	18.899	0.577	17.747	20.051
	2	19.644	0.524	18.597	20.690
	3	18.285	0.567	17.152	19.417

[a] Significant interaction effect.

Table A.10.
Descriptive data of the interaction between game order and gesture

	Game order	Mean	SE	95% Confidence interval	
				Lower bound	Upper bound
Errorsum[a]					
gesture	1	1.233	0.133	0.968	1.498
	2	0.851	0.111	0.629	1.073
	3	0.675	0.096	0.483	0.867
no-gesture	1	0.771	0.133	0.506	1.036
	2	0.790	0.111	0.568	1.013
	3	0.600	0.096	0.408	0.793

[a] Significant interaction effect.

Table A.11.
Descriptive data of the interaction between game order and gender

	Game order	Mean	SE	95% Confidence interval	
				Lower bound	Upper bound
Maxofbeats[a]					
male	1	6.778	0.453	5.872	7.684
	2	6.926	0.607	5.714	8.138
	3	9.148	0.808	7.535	10.762
female	1	5.333	0.377	4.580	6.087
	2	6.051	0.505	5.043	7.060
	3	5.949	0.672	4.606	7.291
Avgofbeats[a]					
male	1	3.681	0.178	3.326	4.036
	2	3.624	0.251	3.122	4.125
	3	4.454	0.316	3.823	5.086
female	1	3.197	0.148	2.901	3.492
	2	3.358	0.209	2.941	3.776
	3	3.386	0.263	2.861	3.912
Non-zeroturns[a]					
male	1	17.926	0.887	16.155	19.697
	2	19.852	0.805	18.243	21.461
	3	17.519	0.872	15.777	19.260
female	1	19.872	0.738	18.398	21.345
	2	19.436	0.670	18.097	20.775
	3	19.051	0.725	17.603	20.500

[a] Significant interaction effect.

About the Authors

Hatice Kose-Bagci received her MS and PhD degrees from the Computer Engineering Department, Bogazici University, Turkey, in 2000 and 2006, respectively. During her MSc and PhD studies, she worked in several research projects involving vision, localization and multi-agent planning in robot soccer, as well as working as a Teaching Assistant and Instructor in Bogazici University. She is currently a Research Fellow at the University of Hertfordshire, working in the EU sixth Framework Project RobotCub. Her current research focuses on gesture communication and imitation in child-sized humanoid robots. Her research interests include autonomous mobile robots, social robotics, interaction, communication and imitation in artificial systems and robotics.

Ester Ferrari is a Research Fellow in the School of Computer Science at the University of Hertfordshire. She is involved in the European IROMEC (Interactive Robotic Social Mediators as Companions) project that investigates the potential use of an interactive, autonomous robotic toy in therapy and education for children with special needs. She completed her PhD in Ergonomics at the Department of Psychology, University of Torino, Italy, in 2007. In the past few years she has conducted research in the areas of human–robot interaction, human-computer interaction, UCD design, usability, ergonomics and psychology.

Kerstin Dautenhahn is a Research Professor in the School of Computer Science at the University of Hertfordshire, where she coordinates the Adaptive Systems Research Group. She received her PhD degree from the Biological Cybernetics Department, University of Bielefeld, Bielefeld, Germany, in 1993, and subsequently worked at GMD (now Fraunhofer) in St Augustin, Germany, the VUB-AI Lab in Brussels, Belgium, and the University of Reading, UK, before she joined the University of Hertfordshire, in 2000. She has pioneered research in social robotics, human–robot interaction and assistive technology, and has published more than 200 research articles. She has edited several books and frequently gives invited keynote lectures at international meetings. She regularly organizes conferences and has been Principal Investigator of her research team in several European projects on developmental robotics (RobotCub), robot companions (Cogniron and LIREC), educational virtual environments (eCircus), and robotics and assistive technology (IROMEC, RoboSkin). She is Editor-in-Chief of the journal *Interaction Studies: Social Behavior and Communication in Biological and Artificial Systems*, as well as an Associate Editor of *Adaptive Behavior*, *International Journal of Social Robotics* and *IEEE Transactions on Autonomous Mental Development*. She is a Fellow of the Royal Society for the Encouragement of Arts, Manufactures and Commerce (RSA).

Dag Sverre Syrdal received his BS in Psychology from Queen's University Belfast, in 2001, and his MSc in Research Methods and Data Analysis in Psychology from the University of Hertfordshire, in 2002. After working as an ESL teacher in China, he joined the Adaptive Systems Research Group, in 2006 where his work has centered around the planning, execution and analysis of user studies within the Cogniron project. His research interests include human–robot interaction (HRI), with an emphasis on socially acceptable interactions in human-centered environments as well as the role of individual differences in HRI, visual attention and the psychology of religion.

Chrystopher L. Nehaniv (BSc, University of Michigan, 1987; PhD, University of California at Berkeley, 1992) is a Research Professor of Mathematical & Evolutionary Computer Sciences at the University of Hertfordshire, where he founded the Adaptive Systems Research Group together with Kerstin Dautenhahn, in 2000. His main interests include interactive systems, complex adaptive systems, artificial life, emergence of language and communication in embodied robots interacting with humans, as well as abstract algebra (theory of semigroups groups and sequential machines) and its applications (especially algebraic biology). He is also an Associate Editor of the journals *BioSystems* and *Interaction Studies*.