Drumming with a Humanoid Robot: Lessons Learnt from Designing and Analysing Human-Robot Interaction Studies

Hatice Kose-Bagci, Kerstin Dautenhahn, and Chrystopher L. Nehaniv

Adaptive Systems Research Group
University of Hertfordshire, School of Computer Science
Hatfield, Herts, AL 10 9AB, U.K.
{H.Kose-Bagci,K.Dautenhahn,C.L.Nehaniv}@herts.ac.uk

Abstract

We summarize methodological and experimental design issues related to three human-robot interaction studies investigating a drumming experience with Kaspar, a humanoid child-sized robot, and (in total 116) human participants. Our aim[1] is not to have Kaspar just replicate the human's drumming but to engage in a 'social manner' in a call and response turn-taking interaction. This requires the set up of enjoyable as well as (as much as possible) controlled experiments. Two Human-Robot Interaction (HRI) experiments with adult participants and one experiment with primary school children were carried out to investigate different aspects of such interactions. We briefly summarize issues concerning experimental methodology and design, as well as ethical, legal, safety issues in addition to many 'practical' challenges of setting up and conducting HRI experiments with an autonomous humanoid robot.

Introduction

We present methodological and experimental design issues related to three exploratory studies investigating a drumming experience (*drum-mate*) with Kaspar [Blow, et al., 2006] and human participants. This research is part of a project in developmental robotics with a particular emphasis of our work on gesture communication. The primary goal of this work is to achieve *(non-verbal) gesture communication* between child-like humanoid robots and human beings, whereby drumming served as a test bed to study key aspects such as turn-taking and non-verbal gestures.

In the first study turn-taking is deterministic and head gestures of the robot accompany its drumming to assess the impact of non-verbal gestures on the interaction [Kose-Bagci, et al., 2007]. The second study focuses on emergent turn-taking dynamics; here our aim is to have turn-taking and role switching which is not deterministic but emerging from the social interaction between the human and the humanoid [Kose-Bagci, et al., 2008]. Each of these two experiments were carried by 24 adult participants (in total 48 adult participants were involved). The third study with 68 primary school children focuses on the effect of embodiment and gestures on the subjective and objective evaluations of the human participants (details of the study and results will be published in a future publication [Kose-Bagci, et al. in preparation]). In all three studies (whose detailed results are reported elsewhere), participants did not have any prior experience with robots. All the experiments were carried out in real-time, and the humanoid robot was operating completely autonomously.

The remainder of this paper is organized as follows, the next section overviews the *drum-mate* studies, their methodology and the research questions motivating them. The experiment design section describes the experimental setups, provides brief information about the humanoid robot Kaspar, and the game setup, followed by a section on data collection. Legal and safety issues, as well as ethical, experimental and other methodological issues are discussed in the following sections. The last section includes a brief conclusion on the experiments, lessons learnt, and presents ideas for future work.

Drum-mate

Methodology

Drum-mate is an interactive drumming game played by a human participant and an autonomous humanoid robot.

[1] 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.

The game was enriched by non-verbal gestures, or different computational turn-taking models according to the research interests of the different studies. The human participant starts the game by playing a simple rhythm on his/her toy drum. When the robot 'understands' that the human has finished playing, the robot takes its turn, trying to imitate the human's drumming. Then the human takes his/her turn at drumming, and so on. This continues for a fixed amount of time, e.g. 3 minutes in the experiments involving adult participants. The robot uses audio feedback to regulate turn-taking and imitation.

Research Questions and Expectations

In the first experiment, we studied the effect of the robot's social gestures in a game of imitation and (deterministic) turn-taking. We expected that participants would be more engaged (in terms of drumming performance) and evaluate the interactions (questionnaires) more positively in the experimental condition when Kaspar used head gestures while imitating the human's drumming than when no such gestures were used [Kose-Bagci, et al., 2007].

The second experimental study investigated the effect of three different probabilistic computational models on the emergent turn-taking dynamics in a drumming game. The game was a modified version of the *drum-mate* game where Kaspar used no gestures but only drumming, and the game mainly focused not on imitation but on turn-taking dynamics emerging from the social interaction between the humanoid and the human participant. Here we expected the different probabilistic controllers to impact the interaction experience significantly. As in the first experiment, objective measures of drumming performance, as well as the subjective evaluations by the participants were analysed [Kose-Bagci, et al., 2008].

The third set of experiments mainly focused on the effect of different embodiments and non-verbal gesture conditions on the interaction between children and the humanoid robot. Like the first two experimental studies with adults, in this work, we also analyse the results in terms of performances of the robot and the human participants and subjective evaluations (questionnaires). Our research interests mainly focused on the differences between conditions where children play in real-time interaction with either the physical robot, the projection of the remotely located robot, or with the 'disembodied' robot (only the sound of the hidden robot is available to the children). Also we expected that these differences would increase in the presence of additional robot gesturing.

Experimental Design

Kaspar

The experiments were carried out with the child-like humanoid robot Kaspar which was designed and built by the members of the Adaptive Systems Research Group at the University of Hertfordshire to study human-robot interactions with a minimal set of expressive robot features. Kaspar has 6 degrees of freedom in the head and neck, 2 in the eyes that are fitted with video cameras, a mouth capable of opening and smiling, and 4 in the each arm. The face is a silicon-rubber mask, which is supported by an aluminium frame [Blow, et al., 2006]. It has immobile legs and fixed feet and hands.

Experimental Setup

The first two experiments (with adult participants) were carried out in a separate room isolated from other people and noise which could affect the drumming experiment. Kaspar was seated on a table with the drum on its lap (Figure 1). The human partner was seated in front of the robot using another drum that was placed on the table. The human participants were to use a pencil to hit the drum. [Although we suggested to the participants to use one pencil and hit the top of the drum, sometimes they used two pencils, or they used their bare hands (single hand or with both hands) and several times they used the tambourine-style bells around the drum's sides.]

The third experiment (with children) was carried out in two almost identical cubicles isolated from the rest of a room with high barriers (Figure 2). In the rest of the room other robotic activities took place at the same time with other children. In one cubicle we had Kaspar seated on a table with the drum on its lap, similar to the first two experiments. In the second cubicle instead of the table and the robot Kaspar, we projected a real time image of Kaspar on a whiteboard, in order to study the effect of the robot's embodiment (physically present versus remotely located). For the third condition, where only the sound of the robot was heard (disembodied robot), we use the same setup but turned off the projector. (See Figure 3). All other aspects of the experimental setup were kept the same.

Fig. 1 A screen shot from the experiments

(a) Physically embodied robot, Kaspar and child participant

(b) Child participant is playing the game in the *projection condition* where the robot's live performance is captured by a camera and projected to a wall in front of the child.

(c) Child participant playing the game with the 'disembodied' robot in the *sound only* condition, where the robot is hidden but audible and only its drumming sound is heard.

Fig. 2 Screenshots from the experiment with children where they played (a) with the physical robot, (b) watching a projected image of the robot, and (c) hearing only the sound of the robot

Interaction Game Setup

For the first two experiments with adults, before every experiment, for each participant, we used a one minute demo of the robot where participants were shown how to interact with Kaspar. Here the participant played the game following a brief introduction of the robot and the game from the experimenter. They learned the rules of the game and got used to the robot without being video recorded. This was followed by three games reflecting the three experimental conditions [Kose-Bagci, et al., 2007; Kose-Bagci, et al., 2008] each lasting three minutes, without indicating to the participants anything about the differences between the conditions.

For the third experiment, we again had three games with different conditions, each of which took two minutes. We had a 30 seconds demo of the first condition in that session which was carried out by one of the experimenters with necessary explanations which was same for all the sessions. So if the first condition is the one with the projected image of Kaspar, then the children see exactly the demo of this condition, not the demo with the physical Kaspar itself. Unlike the first two experiments, the demo is given to a group of children. After the demo the participants will play individually with the robot in the actual experiment.

For each experiment, we used all possible different presentation orders of the games, to analyze the effect of the order of the games on the humans. This is important for avoiding possible fatigue or learning by the participants.

Compared to the experiments with adults, in the experiments with children, simpler gestures were used, and the game duration was decreased to two minutes from three minutes. Also the time between turns was decreased to adapt the game better for the children (i.e. to make the game faster and easy to understand).

Data collection

In the experiments, data were collected to analyse how the human participants evaluate different games and both the robot's and the human participants' performances during these games. We had three main sources of data in our experiments: questionnaires, the drumming data recorded by the robot itself, and the video recordings of the trials including the human partners' drumming behaviour which were then annotated and then quantitatively analysed.

Questionnaires, Consent, Ethics

Before starting the trials, each adult participant was given a questionnaire and a consent and demographics form involving a short description of the experiment and related work. As described in detail in the following subsection, video recordings are important data sources for our studies, so in the consent forms we ask the participants' permission to record their performances by video cameras during the experiments and to use these recordings to produce photos and movies for scientific presentations. We also used these video recordings as data sources to analyse the performances of the participants and the robot. Therefore, participants were also given the option to consent to the video recording and analysis, but not to the use of videos for scientific presentations. If participants do not consent at all to be video recorded during the experiment then their data have to be excluded from the trial.

Unlike some other HRI experiments, our experiments were totally volunteer based, our participants were not paid for their attendance of the trials. This made it difficult to find a large number of participants, with different features e.g. female or left handed participants. Also in the consent forms we inform them that participation in the study is voluntary and that they can leave the experiment at any point during the experiment without being questioned.

Moreover, an ethics approval form including very detailed information about the experiments regarding safety, data collection etc. had to be submitted to the faculty ethics committee of the university. Approval had to be granted before the recruitment of participants and the actual experiments could start. In the case of child participants, in addition a parental consent form was sent to each parent of the children involving detailed information about the experiment and the presence of robots in the experiment, including the possibility of recording the sessions with video cameras and using these recordings later in scientific presentations. According to the result of these forms, some sessions were not recorded by video, or the recorded video was just used for data analysis but not in scientific presentations.

In the first two experiments with adults, the questionnaire which was given before the experiments included general questions about the adult participants, e.g. name, age, nationality, their profession, and if they are parents/careers of children (to understand if they are used to playing with children/children's games).

The children were asked different questions e.g. regarding their tendency to play video games. We were very careful about not asking questions regarding their nationality or ethnic origin, which might be interesting in scientific terms (for cross-cultural comparisons) but may offend or cause discomfort to them. Also we tried to put the questions as simply and understandably as possible, and used small pilot groups to test the usability and understandability of the questionnaires before the real experiments.
After each game each participant completed one page of the questionnaire to express her/his opinion about the game s/he had just played (evaluation of the game and the robot's behaviour in that particular game i.e. sociable/unsociable, or enjoyable/not enjoyable, quality of the interaction for child participants, and the evaluation of the game, drumming of the robot and the social interaction with the robot in the case of adult participants).

Once the participants had completed the items related to the last (3^{rd}) game, they completed the last session of the questionnaire where they could judge the overall experience by deciding which of the three games they liked the best and which they liked the least and the reasons behind that decision.

In adults they also had to judge whether there were any differences between the three games and state these differences.

We tried to keep the number of questions very brief and used simple and direct phrases in the questionnaire.

All the questions about the evaluations were scale based, and the participants were encouraged to write down and express their detailed feelings and suggestions after each evaluation. All of our questionnaires were designed with the help of our psychologist team members for the benefit of the participants.

Video

The experiments were recorded by at least two different cameras positioned at different locations of the experimental area (one facing the human participant and another facing the robot), at each single game. Usually a second experimenter was present in the experimental area to control the cameras, start/stop them during the trials, and (in order not to waste video tape) not to record when the participants played demo games, or worked on the questionnaires between the trials. We used cameras with tapes and fixed them on tripods in several locations of the experiment room where we can view the experiment but do not interfere with the robot or the participant. We even used small tripods to fix the cameras on top of book shelves or doors to get the best viewing angle (to see both faces of the robot and human and the drums).

The video recordings were then analysed manually to detect the performance of the human participant's behavioral data (e.g. the number of drum bouts played by the human, and number of turns taken by the human at each game). Also the video recordings gave clues about the humans' behaviours at certain situation, how they reacted physically and emotionally in different conditions. These observations and the data gathered were compared with the questionnaire results and the data recorded by the robot itself.

Robot

The behavioural data taken from the robot itself includes clues about the robot's and the human participant's performance. Kaspar records its performance (e.g. the number of drum bouts played by Kaspar, and the number of turns taken by the Kaspar at each game). Kaspar also records the human's performance as well (e.g. the duration between two drum bouts) to imitate their performance within its physical limitations. There are two main sources of behavioral data: the audio feedback taken by microphones and analysed in real-time by the robot to extract features of the drumming (i.e. number of drum bouts played, timings and the durations between drum bouts), and the robot's own actions (i.e. name of the action taken at that turn, joint values that are activated during the actions, and timings of the actions).

Note that Kaspar does not have a memory unit on board so this data is recorded in real time on a laptop, not on Kaspar itself .

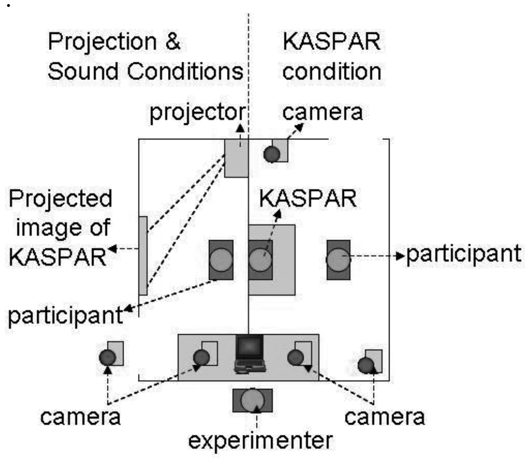

Fig. 3 The experimental layout of the experiment with children playing drumming games with Kaspar in different conditions that varied in terms of the robot's embodiment.

Note, the robot's gestures were kept very simple (e.g. simple head moves, nodding and blinking in the experiments with adults, and additionally a smile to show 'happiness' and a neutral smile to show 'sadness' when the robot could not get any feedback from the human, and waving a hand to say the game is over, for the children participants). More 'expressive' gestures might have distracted participants and could have interfered with our research agenda.

Legal and safety issues

In our experiments we worked with only one humanoid**,** Kaspar, which was risky in case the robot broke down. Its power supplies were low voltage (6V and 12V) rechargeable lead-acid gel batteries for safety. The batteries should be charged fully before each trial, especially if the robot was very active during the experiments. Therefore we limited the use of the robot to a few hours a day. We always had access to a researcher responsible from the maintenance of the robot. The robot never interacted physically with the human, e.g. by touching, and it was stable on the table and did not move its body during the trials, except its head and arms. Kaspar was placed on a table and we always kept a safety distance (at least 30 cm) between the robot and the participants.

All of our researchers who worked with children had CRB clearance (CRB- Criminal Records Bureau, UK). During the event involving children visiting our University (with an opportunity of participants to experience various demonstrations of interactive software and robots, see [FearNot! event, 2008]), their teacher, and a psychologist from our team accompanied the children. The questionnaires were prepared and given to the children by the psychologist. Neither adults nor children were left alone with the robot. The children were not allowed to have sweets during the whole event, but had lunch which was prepared according to their dietary requirements.

Ethical Issues

Although our experiments targeted specific research questions they were also designed to be 'fun' for the participants who volunteered to take part in the study. This was particularly important for the trials involving children who visited our university as part of a school excursion that was meant to have an educational but also enjoyable nature. Thus, we had to use a limited number of experimental conditions in order not the make the event boring for them. For the same reason the duration of the experiments had to be kept short, and overall the experiments had to be designed in a 'pleasant' manner (participants should feel relaxed and comfortable during the experiment). Our intention was to create a playful and engaging setup for the study of human-robot interaction, so many aspects of the work could not be as rigorously controlled as may have been desirable from a purely experimental design point of view. These ethical issues are important to consider in particular for children and other vulnerable people, since researchers do not want to "waste the time" of participants who volunteer in our studies and thus contribute to our research.

Experimental and other methodological issues

Before the experiments are set up, the experimental area and all equipment involved needs to be checked carefully, e.g. check that the video cameras and robots are working properly etc. Also, once the experimental design and setup have been decided, running a simulation of the experiment (whereby experimenters may take the role of participants) is important to see if the time restrictions are satisfied. This is a must especially when running experiments with lots of participants on the same day. Especially in the experiment with children, we had to consider the timing for each child, including their preparation (entering the experimental area, sitting down etc.), the duration of the actual interaction experiment and the time required to complete the questionnaires. In addition, the experimenters need to prepare and practice explanations and answers to possible questions from the children about the work, including questions about the robot's functioning etc. Note, this requires to a) provide as much detail as necessary in order to satisfy the curiosity of the children, and b) not to disclose too many details that may confuse or overwhelm the children, or introduce strong biases influencing the outcome of the experiments.

We also decided to work with the children in teams, not alone, which would help them to get used to the robot and the game. But working with teams of children lead to other issues, like how to identify the children individually (and match the ID codes that they were given to the results of the data collected during the experiment etc.).

In addition to within-subject comparisons (e.g. for testing different game conditions and different gestures of the robot) we also carried out between-subjects tests in order to study the impact of individual features of the participants which may affect the results, e.g. gender.

Our experiments were 'controlled' (in the sense that each participant in a particular study was exposed to a specific and clearly defined experimental set up whereby a particular experimental procedure was followed involving different experimental conditions) but they were not fully restricted laboratory experiments because we intended to create a 'playful atmosphere' in order to facilitate natural interaction that is emerging between the robot and the human and any adaptation of the human to the robot and the game. Also, all our participants were volunteers so we needed to achieve an enjoyable task as well as a controlled experiment. Last but not least, the group of children we worked with was an opportunity sample, and while we could control for age (due to the fact that the children arrived as part of a school class) we could not control e.g. for gender.

Audio analysis was a very vital part of our game, and the robot's "hearing" was effected by both external and internal noise (coming from the robot's motors) so we had to use some noise filters or cover our microphones especially in the experiment with the children where we had 10-20 children shouting and talking in the same room (the experimental area was only separated by screens). Note, in the experiments with the children we deliberately decided to carry out the experiments in the same room were the other robotic group activities took place. Alternatively, we had considered isolating individual children or small groups of children and leading them to a different nearby experimental room. However, we decided against such a more 'controlled' approach since it would have interfered with the enjoyable nature of the event (as part of a school excursion in order to learn, playfully, about robots and virtual characters). It also ethically did not seem to be justified to remove individual children from their peer group in this group-oriented event.

In terms of the experimental equipment, we tried to have a "natural" looking robot, using clothes for the robot with neutral and not too bright colors. We tried to keep the experimental area 'tidy', i.e. tried not leave unrelated objects in the experimental area that could distract the attention from the main focus of the experiment.

Demos and explanations to participants were very important in our experiments; they were kept the same for all the participants. Giving slightly more explanation or a smile of the experimenter during the experiment might affect the evaluations of the participants. Also it is important to use the same experimenter for the all experiments since his/her behaviour or characteristics (gender, height, age, tone of voice etc.) may have an impact on participants in the experiments. Ideally experiments could be repeated with different experimenters, in order to reveal any effects this may cause, but practically this usually goes beyond the scope of an HRI study. In our HRI studies the experimenters had a 'passive role' during the experiments, i.e. they were present but did not proactively engage with the participants. The main role was to provide demos, explanations, and generally guide the participants through the experiment.

It is important to use small pilot groups to test the usability and understandability of the questionnaires before the real experiments. Especially when conducting large scale experiments, with strict time restrictions, this is a vital issue. In our experiment with children we could only use one child to test the questionnaire. Ideally several children should be used to get feedback in particular on the questionnaire design.

Surveys and questionnaires might not always provide the 'full picture'. For example, females tend to have higher agreeableness scores than males and participants with higher agreeableness can thus be expected to rate the robot's capabilities as better [Costa, et al., 2001]. Also if some of the participants who knew the experimenter, they might have be less 'subjective', or the primary school children might answer the questions in the same spirit as doing 'homework', and might not express their 'natural' feelings. So it is vital to collect additional behavioral data from different sources e.g. video recording, using different sensors, and comparing the behavioral data and questionnaire data for different aspects. Interestingly, our research [Kose-Bagci, et al., 2007; Kose-Bagci, et al., 2008] showed that results from different data sources might not ‘agree’. For example, when judging the most preferred game, and least preferred game, although a participant might dislike the drumming of robot in one game and evaluate the game in the questionnaire according to this, he/she might have scored the least error in terms of objective measures of drumming performance in that game

It is important not to explain fully the overall experimental design and the game, not even details of the robot (in the embodiment experiments) to the participants before the start of the interaction experiments e.g. in instructions because the aim is to observe the adaptation of the human to the robot and different aspects of (non-verbal) communication emerging from the social interaction between the human participant and the humanoid. If the participant starts with too much information this may bias him/her and affect the results. Sometimes participants spend some time to explore the robot rather then playing the game, which increases performance error according to objective behavioural measures, but we did not intervene since we considered it a part of the interaction. We used the demo session to address this issue. Still participants asked questions at the end, for example: “Does the robot learn?”, “Can it see me?”, “Why does not it talk?” One of the participants claimed that the robot smiled at him ‘badly’ when he did something ‘wrong’ although the robot was not making any facial gestures at all.

The duration of the experiments is important. Our experiments with adults lasted 3 minutes and those with children 2 minutes. The experiment should be long enough to collect an adequate amount of data and short enough not to be boring since boredom also effects the evaluations of the participants. Even if the task is enjoyable and the humanoid is interesting, in experiments involving a repetition of movements or tasks, doing the same thing for several minutes is not always pleasant.

Although we worked with groups of participants at different times (these three experiments were completed over more than 1.5 years), we kept all the experimental conditions identical (e.g. the experimental setup and the robot gestures used), within the same experiment, which was a hard task when using a robot (which had been used extensively during the same period for several other research projects), and dealing with more than 100 human participants.

In all of our experiments, the robot was autonomous. Therefore it was important to avoid the belief in participants that the robot was remotely controlled by a Wizard-of-Oz (WoZ) technique, see e.g. [Green, et al., 2004]. During the trials we tried to avoid using the control laptop, because when the experimenter worked with the laptop, the participants might have thought that she was controlling the robot. Moreover, in the embodiment experiments, the children could think that not the robot but the experimenter was playing the drum in the disembodied condition (when the robot was hidden but its drumming sound was heard), so we always kept the experimenter in view of the participants, but not watching the participants, not using the laptop, and not interfering with the participant or the robot. Rather, the experimenter did something seemingly ‘irrelevant’ to the study, i.e. reading a book. Being watched may put stress on the participants.

The selection of the robot that was used in the HRI experiments is also an important issue. Some humanoid robot’s are functional and robust from the experimenter’s and designer’s point of view, but might look ‘scary’ especially from children’s point of view, compare research on the ‘uncanny valley’, e.g. [MacDorman & Ishiguro, 2005]. Some participants find the inner noise of the robot operating ‘normal’ as this makes it more ‘robot-like’, but it can be annoying for others.

Some people may have concerns towards robots which may prevent them from interacting with the robot ‘naturally’. Such participants would tend to behave in a manner less ‘relaxed’ and ‘open’ towards the robot. Such an attitude might be hard to recognize in questionnaire data but can be detected in behavioral data. It is essential to gauge people’s feelings regarding and attitudes towards robots in order to detect participants with strong negative feelings towards robots. Generally, participants’ personality profiles, individual interests, hobbies etc, may also provide useful data that may explain how people react to and interact with robots in HRI experiments.

Related work

Other researchers have identified various important methodological issues in HRI research. However, a full survey of related work goes beyond the scope of this paper. Illustrating related work, Walters and colleagues (2005) also have provided a very useful discussion of the practical and methodological aspects of HRI studies which were based on several HRI experiments. They describe the legal and safety issues in detail. Those experiments took a human-centered perspective in HRI studies with a human-scaled mobile robot which was primarily controlled by the WoZ technique. The methodological issues related to these experiments, which are slightly different from those of our experiments, were described in detail. Table 1 shows a comparison of both works. Importantly, while Walters et al. (2005) used a primarily remote controlled robot, our experiments have taken a dual perspective: developing autonomous behaviours for a humanoid robot to play interaction games with people, while at the same time assessing the behaviour of people playing interaction games with the robot and their subjective evaluations of the games.

Table 1. Comparison of Drum-mate and Walters et al. Wizard-of-Oz (2005) studies

	Drum-mate studies	[Walters, et al., 2005] studies
Robot platform	Kaspar, a child sized humanoid	PeopleBot™, a human-sized wheeled robot base, extended
Mobility of the robot	Immobile, just head and arms move	Mobile
Control of the robot	autonomous	WoZ + (autonomously in a small scale)
Appearance	Human-like features	Mechanical looking
Experiments with participants	One participant at a time	One participant at a time and groups of participants
Experimenter	In view of participants to prove the robot is operating autonomously	Experimenters controlling the robot are hidden from the robot, the experimenter introducing the participants to the experiment etc. is in view of participants
Perspective of the trials	Both robot and human centered (development of interactive games for a humanoid robot, as well as the study of people's behaviour and subjective evaluation in interaction experiments)	Human centered (studying perceptions, attitudes and behaviour of people towards robots)

Conclusion

We presented the experimental design and the related issues result of three interaction experiments with an autonomous humanoid robot, involving in total 116 human participants playing human-humanoid drumming games. Despite the issues related to the experimental environment, participants, and the robot itself, we had significant results in terms of non-verbal and timing aspects of interaction, imitation, turn-taking and gender differences that are reported elsewhere [Kose-Bagci et al,. 2007; 2008; in preparation].

The methodological issues, and approaches taken to address these and other issues presented in this paper will inform future studies related to human-humanoid social interaction. Many of these issues will also play a role in other HRI experiments, including different application areas such as entertainment, service robots, and educational/therapy robots. Thus, we hope that this paper will be useful for other HRI researchers, in particular those with no or little experience in carrying out user studies.

References

Blow, M.P., K. Dautenhahn, A. Appleby, C. Nehaniv, D. Lee, "Perception of robot smiles and dimensions for human-robot interaction design*",* In *Proc. IEEE RO-MAN06,* pp.469-474, 2006.

Costa, PT; Terracciano, A; McCrae, RR. Gender differences in Personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*. 81:322–331, 2001.

FearNot! event, http://www.herts.ac.uk/news-and-events/latest-news/Antibullying-software-and-robots.cfm, Last accessed: 22 December, 2008.

Green, A., Huttenrauch, H., K. Severinson-Eklundh, "Applying the Wizard-of-Oz framework to cooperative service discovery and configuration", In *Proc. 13th IEEE international workshop on robot and human interactive communication (RoMan2004), Kurashiki, Japan*, 575–580, 2004.

Kose-Bagci, H., K. Dautenhahn, D. S. Syrdal, and C. L. Nehaniv, "Drum-mate: A Human-Humanoid Drumming Experience"*, IEEE-RAS Humanoids2007*, 2007.

Kose-Bagci, H., K. Dautenhahn, and C. L. Nehaniv, "Emergent Turn-Taking Dynamics in Drumming Games with a Humanoid Robot" , *Proc. IEEE RO-MAN 2008*, 1-3 Aug. 2008, Technische Universitat Munchen, Munich, Germany, pp. 346-353, 2008.

MacDorman, K. F. & Ishiguro, H., "The uncanny advantage of using androids in cognitive science research", *Interaction Studies, 7*(3), 297-337, 2006.

Walters, M. L., S Woods, K L Koay, & K Dautenhahn, Practical and methodological challenges in designing and conducting interaction studies with human subjects. *Proc. AISB'05 Symposium on Robot Companions Hard Problems and Open Challenges in Human-Robot Interaction*, 14-15 April 2005, University of Hertfordshire, UK, pp. 110-120, 2005.