Theatre as a Discussion Tool in Human-Robot Interaction Experiments.

A Pilot Study

A. R. Chatley, K. Dautenhahn, M. L. Walters, D. S. Syrdal, B. Christianson
School of Computer Science
University of Hertfordshire
Hatfield, Hertfordshire, England
{a.r.chatley, k.dautenhahn, m.l.walters, d.s.syrdal, b.christianson}@herts.ac.uk.

Abstract— **In the field of Human-Robot Interaction (HRI), a novel experimental methodology is presented for carrying out studies which uses a theatrical presentation with an actor interacting and cooperating with robots in realistic scenarios before an audience. This methodology has been inspired by previous research in Human-Computer Interaction. The actor also stays in role for a post-theatre session, answering questions and encouraging the audience to discuss their respective opinions and viewpoints relating to the HRI scenario enactment. The development and running of a first exploratory pilot experiment using the new Theatre HRI (THRI) methodology is presented and critically reviewed. Based on this review and the associated findings from the audience discussion session, it is concluded that the Theatre-based HRI (THRI) methodology is viable for performing HRI user studies.**

Keywords; Theatre, Human-Robot Interaction, User Evaluation, Robot Performance, Usability, Scenarios.

I. INTRODUCTION

Live Human-Robot Interaction (HRI) user studies, in particular those involving medium or large human-scaled mobile robots, are usually time consuming and typically involve significant resources [1,2,3,4]. Due to practical limitations, participant numbers are usually relatively low, typically less than 20 with few studies involving more than 50 users. To perform HRI studies with larger numbers of participants, Video-based HRI (VHRI) methods have recently been proposed and used as a useful means to obtain early user feedback [5,6,7]. However, video-based experiments cannot replace live experiments [8]. In order to combine live HRI experiences within a more efficient experimental group setting, we take inspiration from the field of Human-Computer Interaction (HCI) where theatre has been used to gain feedback from potential user audiences [9]. Forum style group discussions are commonly held afterwards, with the actors in attendance, to help facilitate open and free communication of thoughts and opinions. Our assumption was that similar experimental methods applied to HRI will prove effective in gauging audiences' opinions regarding robots early in the design development process. The findings would also guide the development of future VHRI and live trials. Using an actor in the role of a robot owner throughout the trial and subsequent discussions can also overcome possible influences on participants due to having the robots' creators present.

This research is associated to the interdisciplinary European FP7 Integrated project LIREC (LIving with Robots and intEractive Companions [10]) which develops and evaluates new companion technology. Within the project, the University of Hertfordshire LIREC research team focuses on the notion of a robotic home companion and is very experienced in using new media forms such as video in their experiments [5,6,7]. This familiarity helps to bridge the gap between disciplines when having actors working with robots. The exploratory experiment presented here draws upon our previous work, which used VHRI experimental methods, to ascertain what type of scenarios might be applicable to the theatre medium. Potentially, findings from Theatre-based Human-Robot Interaction (THRI) studies can be used to further enhance HRI scenarios, to develop and supplement future THRI, VHRI and live HRI experiments, and also contribute to further novel HRI methodology developments [11].

The THRI methodology is based on similar uses of theatre in HCI user studies, pioneered by Alan Newell [12]. The differences between the use of theatre in HCI studies and our use for HRI studies are primarily due to the artifacts involved (i.e. the robots, the subject of the experimental investigation) being used as actors [13]. In previous HCI studies, the scripts for the theatre enactment took into account the display limitations of a computer system. This is similar to actors in a typical theatre production working alongside a prop or particular stage device such as lighting and scenery. The scripts for the THCI studies need to be detailed enough so that the actor is able to work within the limitations of the robots.

The theatre audience should become absorbed by the demonstration and they can be comfortably guided through post-enactment discussions by a neutral actor as the main go between for the audience and the design team [14]. Method actors can also improvise to the situation as required [12], though this is not something that can often easily be achieved with pre-scripted robots.

The methods for using theatre in HCI work well on two levels; one is the discussion aspect through the actor, the other is through the revision of scenarios and working through of ideas actually during the discussion. In THRI it is not presently technically possible to have instant revisions, but it is possible for the actor to hold a thought-process demonstration with the audience. The THRI study methodology would be especially useful for initial user trials, controlling the robot directly (Wizard of Oz (WoZ) [1,2] method), before large resources are committed to programming the robots to work autonomously in full live trials. Therefore, the proposed THRI methodology provides freedom of creation to scenario and robot designers, but simultaneously facilitates early feedback from target user groups [3-8].

II. Methodology Development

The methods discussed previously for HCI user studies were adapted for evaluating robots. Direct robot control (WOZ) allows the robots to act as (semi-) actors in scenes and also allows actions to be simulated that are currently not available for the robots to do themselves (e.g. robust navigation in everyday environments, smooth natural language interaction, sophisticated reasoning and sophisticated ‘social intelligence’ in interaction and dialogue with the user). It is therefore possible to look forward into ‘what if’ scenarios. The scripts used were based upon existing scenarios that the UH research team had previously developed for use at the UH Robot House (an experimental but naturalistic environment suitable for HRI studies in a domestic setting, cf. [6,15]). The scenarios implemented were purposefully basic to allow for an easier transition into the new methodology. However, key tasks, such as the role of fetching objects, were maintained so as to showcase interactions the robots can usefully perform. Note, fetch-and-carry behaviors are typical for robot assistance scenarios in a home environment (e.g., COGNIRON [16]). A Pioneer and a PeopleBot robots from Mobile Robots Ltd were used for the study (Fig. 1) and both robots have been modified by the LIREC team at the University of Hertfordshire. The experiment was designed with some key stipulations:

- The role of an Actor is to improvise any problems that might occur with the robots. The Actor also will remain in character after the experiment, with the pretense of owning the robot.
- The role of a Facilitator to help answer any questions that may be posed by the audience that the Actor is unable to answer.
- The role of a Scientific Facilitator, to field any scientific questions that the neutral Facilitator is unable to deal with.
- A wide range in audience backgrounds and age, to test the methodology itself, not the participants, and to give a benchmark on the working process.
- The use of scenarios that show issues in a modular format in order to later (i.e. during the analysis of the experiment) link audiences’ reactions to individual scenes.

Of the above stipulations, the first two are common to previous experiments in HCI [9,12,14]. Utilizing the actor in character after the event helps to make the audience believe that the experiment was a true representation of an owner and robot. Ideally the interaction between the actor and audience should flow smoothly, but the neutral facilitator is there to offset any problems that may arise if this fails. The scientific facilitator exists purely in a technical role and hopefully their presence is not required. They will not be in the audience’s line of sight, but will still remain near the main discussion area. The wide range of audience participants was decided upon to provide variety for the exploratory study, to focus attention onto the methodology, and to reduce specific target audience related problems. The scenarios cover three relevant areas of HRI; Fetch and Carry, Interactions and Migration:

Fetch and Carry describes the robot moving to and from specific positions, responding to commands to deliver food items etc.. *Interactions* are of two types: Verbal Interaction relates to the direct communication (speech) that occurs between the robot and actor. Non-Verbal Interaction pertain to the robot’s expressive behaviors. *Migration* is the movement of a robot’s ‘mind/Artificial Intelligence’ ('personality') from one robotic platform to another. This is relevant to the LIREC research at UH and investigates issues of embodiment and personalization of the robots [4].

Each scenario should last no more than five minutes, which is sufficient time to give an introduction to each of the three topics. The scenarios provide a 'thought tool' to provoke discussion on what is viewed, and findings from these discussions are used to evaluate both the methodology and the scenarios. For future THRI experiments the scenarios could be more in-depth to give more context for the ‘why’ and ‘how’ of what the audience is shown. A temporary stage setup within an existing room provided an area for the audience to view the

Figure 1. Experiment Robots

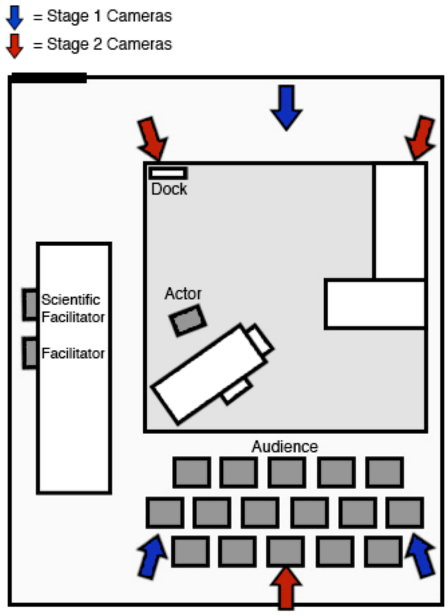

Figure 2. Viewing Area Original

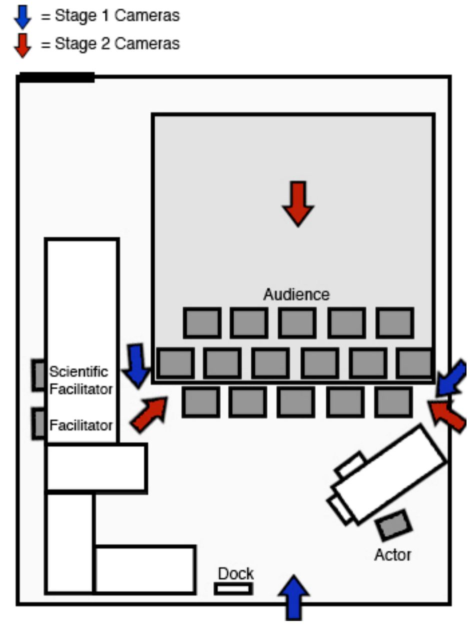

Figure 3. Viewing Area Post-changes

robots in use. Larger THRI experiments would probably use a theatre location with built-in support for lighting and technical facilities. The viewing area was originally planned as shown in Fig. 2, with the layout mimicking a small studio style home: A dining area front left of the stage, a living area at back center, kitchen to the back right, and the entry door at front right. Fitting the three scenario locations into one area was a main goal when planning space requirements. If a larger stage is available then the areas can be moved outwards, but working with a stage does limit the 'safe' area for the robots and humans (i.e. where they can move around safely without being damaged and without hurting any humans being present). Safely for the people and the robots is a crucial aspect to be taken into account.

The audience are in front of the stage and remain there for the follow-up discussion. Off-stage and to the side of the audience is where the robots are controlled from; out of the audience's direct line of sight to allow the audience to 'forget they are there', but where the robot operators can easily see the robots. The robot operators should be able to be contacted easily by the actor.

III. Experiment Implementation

The actual implementation of the experiment required various changes to the original plan. The location for the smaller experiment required a drastic change in the size and organization of the stage. The available space did not provide a 'safe area' for the robots to perform without risk of damage (i.e. falling off the stage). Once the problem was discussed, the stage was altered into a tiered seating area for the audience and the robots on the ground level (Fig. 3). The revised layout provided a much larger robot safe area.

A. Procedure

The participants consisted of an audience of eight, whose ages ranged from early twenties to mid fifties, with various backgrounds including computer science, engineering, humanities, and art and design (media). The audience were lead in and asked to complete a consent form. Next an initial introduction talk was provided by the facilitator. The scenarios lasted for a total duration of 15 minutes and consisted of an introduction by the actor whilst the robots were setup during the beginning of each scenario, followed by the actual demonstration of the robots in role. The first and second scenarios went as planned, using both the Pioneer and PeopleBot respectively. The scenarios themselves proved to be adequate and the actor and robots were able to perform most functions to the level required for discussion. However, Migration proved troublesome technically as the delay in switching control of the robots when under direct control was longer than expected. This problem had not occurred previously in our research and could not be fixed on short notice. This issue meant that the second robot's role (the Peoplebot), was 'acted out' (i.e. simulating the location and movements of the robot) by the facilitator to showcase the required information as best as possible. The audience responded well to the interaction in this manner. All three scenarios were shown, then the discussions were held afterwards. The discussion that followed lasted for the duration of 35 minutes and covered a variety of topics, the facilitator and actor answering questions whilst seated in front of the audience. The entire experiment was filmed from three angles with camera placement changing from two on the robots/actor and one on the audience during the second stage, to two on the audience and one on the facilitator/actor during stage three.

	Avg.(m)	Freq.	Max.(m)	Min.(m)	Total.(m)
Communication	1.25	10	2.5	0.25	12.5
Personalization	1.21	6	1.5	0.25	7.25
Embodiment	1.03	8	2	0.5	8.25
Memory	0.9	5	2.25	0.5	4.5
Intelligence	2.5	1	2.5	2.5	2.5

TABLE 1. Top Level Transcription Results

IV. RESULTS

A. Methodology Evaluation

The discussion following the scenarios was run with more involvement by the facilitator than had originally been assumed. The participants seemed reluctant to start a discussion initially, but once started, showed a willingness to discuss what they had seen with the actor and the facilitator. Intra-audience discussion was, however approached, not achieved to any large degree, apart from some discussions, which took place towards the end of the session. The audience did take the initiative in returning to the discussion topics they found the most salient, without being prompted by the facilitator.

The main aims for the discussion were to primarily gain the audience' opinions on the scenarios and robot appearance and these topics were raised successfully, indicating the viability of the methodology.

B. Scenarios Evaluation and Analyses

Evaluation of the audience responses was initially achieved by breaking down the topic categories discussed in terms of the duration of a discussion segment, along with the number of times a topic was raised. A change of topic was recorded when the topic changed for more than 15 seconds. Both the number of segments as well as their duration was used in the analysis. The frequencies and durations of the different topics can thus be considered a operational measure of what aspects of the scenario that was presented, were the most salient to the audience. This form of data analysis is similar to that suggested in Rosenthal [17].

Table. 1 shows the distribution of frequencies in terms of discussion segments and durations of these segments organized by the five topic areas. The Intelligence category was discussed only once, but for the longest period. The other four categories are closely grouped in terms of overall mean time, for their discussion segments, each approximately 1 minute, but there is a clear distinction between them in terms of frequency. For frequency, Human-robot Communication was raised most frequently. Personalization (i.e. how to personalize technology to individual user's preferences, needs etc.) has a higher number of longer duration discussions, but the topic was

	Avg.(m)	Freq.	Max.(m)	Min.(m)	Total.(m)
HR Verbal	1.69	4	2.5	1.25	6.75
HR Non-Verbal	1.08	3	1.5	0.5	3.25
RH Verbal	0.25	1	0.25	0.25	0.25
RH Non-Verbal	1.13	2	1.75	0.5	2.25

TABLE 2. Communication Discussion Analysis

brought up relatively infrequently. Discussion segments regarding Embodiment (the robot's appearance, movements etc.) has a mean duration of 1.5 minutes with a frequency comparable to the previous two. Lastly, discussion segments involving (robot) Memory has a relatively low frequency, and the mean discussion period is low, suggesting that the audience did not find this aspect of the scenario as noteworthy as the other categories. The frequencies show that, with the exception of Intelligence, the audience returned to each topic repeatedly, suggesting that discussions on one topic were often built on, and frequently led on to, discussion on other topics.

The descriptive analysis of topic categories reported above indicates that the facilitated discussion was wide ranging, with the focus shifting consistently back to the audience' main topics of interest such as communication and embodiment. Communication, leading in both average duration and frequency, was consistently brought back into discussion by the audience meriting further breakdown into subcategories (Fig. 4). Table 2 and Fig. 5 show the breakdown for the top-level topic of Communication into four sub-components; Human-Robot (HR) verbal, HR non-verbal, Robot-Human (RH) verbal, and RH non-verbal. However looking purely at these descriptives does not give an accurate opinion of the audience's attention.

Personalization of the robot (including communication

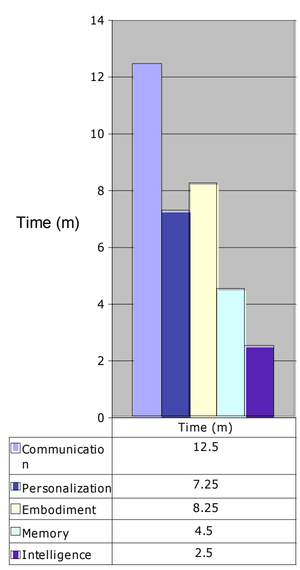

Figure 4. Top Level Transcription Results

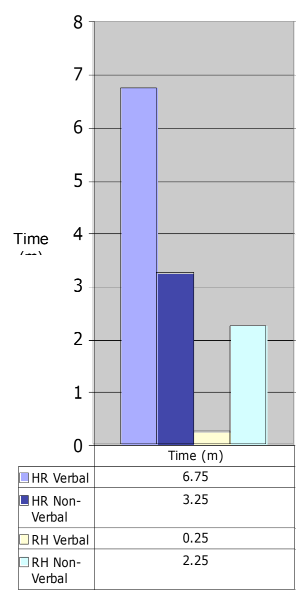

Figure 5. Communication Analysis

personalization) was discussed, often along with memory and embodiment. The personalization of communication was repeatedly brought up. Having the ability to tailor the robot's speech to something that was of their preference seemed important and widely accepted, whilst the methods of communication and responses varied more widely. The link between personalization and communication is of particular interest as it demonstrates how important the issue of communication is in terms of saliency. The audience were split in their preference of communication methods to inform the robot of their needs; hand gestures, vocalized sentences, vocalized phrases, non-verbal movements separate to hand gestures, and vocalized key words. One vocal participant stated "So I would like to see a robot that can respond to more than 'fetch', 'drink', 'now'", thus showing a clear choice in non-keyword communication methods. Other users' expressed views were more ambiguous such as one being "..happy talking to it as long as it was reliable in recognizing what I said".

In terms of Embodiment, Personalization was discussed in its relevance to changeable body parts such as the arms and 'head' (the camera on the pioneer was accepted by the audience as its head). Also personalization related to changing colors and visible appearance to tailor the robot to what the user saw as the ideal embodiment. These discussions showed little preference for a head (Peoplebot) or a camera (Pioneer), however the discussion of usefulness to the individual in terms of personalization was raised and should be taken further in future scenarios to investigate the audience's opinions on customization to specific tasks. The audience did comment that the current implementation was "too cumbersome", a view that was initially stated by the actor during the start of the discussion (as a deliberately chosen strategy to increase the believability of the actor as an actual robot owner). The audience were noted as agreeing with this statement later in the discussion, one participant even stating that in an emergency they would not trust the robot to function and instead would "rather have a dog". Personalization was also brought up in relation to memory and the ability to remember users' schedules. The audience were vocal on how much control they desired over their robots and what tasks they liked them to perform. Remembering tasks and preferences was shown in the scenario by the robot's ability to remember appointments, and the audience thought this was a valid point and worth further exploration.

V. Discussion

The 'Theatre' scenarios used included both verbal and non-verbal communication and as expected this became a main topic of discussion. Robot speech was included since it provided a 'natural' and 'attractive' means of human-robot communication for on-stage. However, the scenarios that are currently used by the research team, and implemented in the Robot House using autonomous robots, show little robot vocalization and most cues are non-verbal. It appears that the use of robot speech in the Theatre scenarios strongly pointed audience's attention towards this topic, while our actual research in HRI on developing autonomously operating robots focuses on a variety of different areas other than speech. Therefore, our future THRI scenarios will attempt to incorporate a variety of (non-verbal) communication methods, more closely matched with current technology. If in future scenarios the robot is clearly using a set method of communication, one that can be explained in a summary before the discussion, then the discussion time utilized for the speech based communication topic could be moved to other areas.

Whilst some ideas of the research team on human-robot communication (e.g. using mobile phones as communication devices – which are not easily visible and detectable by a large audience) cannot be shown on a stage, the adaptation process can work around issues regarding visibility of the robots to the audience. Theatre is not a perfect medium for dispersing information to a wider audience, but it does give a relatively high level of interaction and feedback compared to other HRI user study methods. For example, video-based (VHRI) methods can show robots interacting with a user, but often the visible area is limited. Theatre shows the larger picture, but at the cost of finer details such as close inspection of the robots' actions, sounds levels, and there may be height or visibility issues for smaller robots. An interesting area for future scenario designs and investigations concerns the issue of the actor's willingness to interact with the robot as a 'social being', i.e. the actor gestured, spoke, and used keywords throughout the scenarios, thus giving the impression of the robot responding in the above mentioned manners.

VI. Conclusion

This paper presented a novel experimental methodology for carrying out Human-Robot Interactions (HRI) studies, using a theatrical presentation with an actor interacting with robots in realistic scenarios before an audience. This methodology had previously been used in HCI, but to our knowledge this is the first application to the field of HRI. For the novel THRI methodology, this exploratory study shows that the method of discussion can be effectively lead by the facilitator on stage and via the presented scenarios. Communication was found to be a topic of great interest for the audience in the post-theatre discussions. However, it can be expected that the audience have other topics they are willing to discuss and it is the role of the facilitator and the scenarios to try and balance the discussions. The communication issues highlighted and discussed in this exploratory THRI presentation, link to previous HRI work

[15] and show that the THRI methodology is accurate in its findings regarding initial audience responses. What remains to be tested are the limits of controlling the audience participation and discussion, without influencing their decisions and opinions.

Our research in LIREC on robot companions can take away from these studies: a) in concrete terms, ideas concerning personalization and other topics brought up by the audience, and b) conceptually, relevant scenario information from the discussion in the ideas behind structuring scenarios to focus the audience onto specific topics. A structured list of scenarios that can be transported and shared between robots and the discussion medium (HRI, VHRI, THRI) is an ideal that should be achievable. Future use of this methodology with these changes should further enhance the process of development for scenarios that are impacted by communication, personalization, embodiment, memory and intelligence.

Future development of THRI as a methodology for discussing specific topics relies on orderly changes to the scenarios used, along with the staging areas being given more thought in the initial design process. Theatre has many devices inherent to the use of a staged viewing area, and these also need to be incorporated into the THRI methodology. By giving further thought to the environment in which the methodology is to be used, it will also be possible to plan scenarios to pre-empt the topics of discussion. If the discussion can be pre-planned to some degree, the scenarios can undergo a refinement process prior to being used on the stage – thus fine tuning the discussion to topics related to the aim of the THRI experiment.

The exploratory study shows the need for prior analysis of scenarios before entering into the theatre. The theatre itself may overcome some problems that the exploratory study uncovered: Technical and stage problems can be dealt with more easily with a professional theatre environment, and the discussion problems with a trained facilitator. Future use of the methodology reflecting these goals would prove more effective to a wider genre of HRI and theatre.

Acknowledgment

We would like to acknowledge the actor David Harold for his participation. The work described in this paper was partially conducted within the EU integrated project LIREC (Living with Robots and intEgrated Companions). Funded by the European Commission under FP7-ICT under contract FP7-215554.

References

[1] A. Green, H. Huttenrauch & K. Severinson-Eklundh, “Applying the Wizard of Oz Framework to Co-Operative Service Discovery and Configuration”, Proceedings of the 13th IEEE International Workshop on Robot and Human Interactive Cmmunication (RO-MAN 2004), Kurashiki, Okayama Japan, (2004), pp. 575-580.

[2] A. Green, E. A. Topp & H. Huttenrauch, “Measuring Up as an ntelligent Robot – On the Use of High-Fidelity Simulations for Human-Robot Interaction Research”, Proceedings of Performance Metrics for Intelligent Systems Workshop (PerMIS ’06), Gaithersburg, MD, USA, (2006), pp. 247-252.

[3] M. L. Walters, K. Dautenhahn, S. N. Woods, K. L. Koay, R. te Boekhorst, D. Lee, "Exploratory Studies on Social Spaces between Humans and a Mechanical-looking Robot", Connection Science 18(4) , (2006), pp. 429-442.

[4] D. S. Syrdal, K. Dautenhahn, M. L. Walters, K. L Koay “Sharing Spaces with Robots in a Home Scenario – Anthropomorphic Attributions and their Effect on Proxemic Expectations and Evaluations in a Live HRI Trial”. Proc. AAAI Fall 2008 Symposium ‘AI in Eldercare: New Solutions to Old Problems’ November 7-9, Washington, DC, USA, accepted (2008)

[5] M. L. Walters. “The Design Space for Robot Appearance and Behaviour for Robot Companions”. PhD thesis, University of Hertfordshire, (2008).

[6] S. N. Woods, M. L. Walters, K. L. Koay, K. Dautenhahn, "Methodological Issues in HRI: A Comparison of Live and Video-Based Methods in Robot to Human Approach Direction Trials", Proceedings of 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 06), Hatfield, UK, 6-8 September (2006), University of Hertfordshire. pp 51--58.

[7] S. N. Woods, M. L. Walters, K. L. Koay, K. Dautenhahn, "Comparing human robot interaction scenarios using live and video based methods: Towards a novel methodological approach", Proceedings of the Ninth International Workshop on Advanced Motion Control (ACM 2006), Istanbul, March 27-29, (2006), pp. 750--755.

[8] K. Dautenhahn “Methodology and Themes of Human-Robot Interaction: A Growing Research Field”. International Journal of Advanced Robotic Systems 4(1) (2007), pp. 103-108.

[9] M. Rice, A. F. Newell, M. Morgan, "Forum theatre as a requirements gathering methodology in the design of a home telecommunication system for older adults", *Behaviour and Information Technology*, 26(4) (2007), Taylor and Francis, pp.232-331

[10] LIREC, “Living With Robots and Interactive Companions”, [Online], Available: http://www.lirec.org/ [Accessed: Oct. 12, 2009].

[11] J. Anstey, D. Pape, S. C. Shapiro, O. Telhan, T. Nayak, "Psycho-Drama in VR," Proceedings of The Fourth Conference on Computation Semiotics (COSIGN 2004), University of Split, Croatia, 2004, 5-13.

[12] M. Morgan, A. F. Newell, "Interface between two disciplines - the development of theatre as a research tool", *Proceedings HCI 2007*, 1 (2007), pp.184-193. Beijing, 22-27 July

[13] C. Breazeal, A. Brooks, J. Gray, M. Hancher, J. McBean, W.D. Stiehl, and J. Strickon. "Interactive Robot Theatre", *Communications of the ACM*, 46(7), (2003), pp. 76-85.

[14] A. F. Newell, A. Carmichael, M. Morgan, A. Dickinson, "The use of theatre in requirements gathering and usability studies", *Interacting with Computers*, 18 (2006), Elsevier, pp.996-1011

[15] D. S. Syrdal, K. Dautenhahn, S. N. Woods, M. L. Walters, K. L Koay, “Doing the Right Thing Wrong' - Personality and Tolerance to Uncomfortable Robot Approaches. Proc. The 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN06), University of Hertfordshire, 6-8 September, Hatfield, UK, (2006), pp. 183-188, IEEE Press.

[16] COGNIRON, “The Cognitive Robot Companion”, [Online], Available: http://www.cogniron.org/final/Results.php [Accessed: Oct. 12, 2009].

[17] R. Rosenthal, & R. L. Rosnow,”Essentials of behavioral research: Methods and data analysis”, (3rd ed.). New York, McGraw Hill, (2008), chapter 9.