# Methodological Issues Using a Comfort Level Device in Human-Robot Interactions

Kheng Lee Koay, Michael L. Walters, Kerstin Dautenhahn

*Adaptive Systems Research Group, School of Computer Science*
*University of Hertfordshire*
*College Lane, Hatfield, Herts, AL10 9AB, United Kingdom*

{K.L.Koay,M.L.Walters,K.Dautenhahn}@herts.ac.uk

*Abstract* – **This paper introduces a handheld Comfort Level Device to measure subjects' comfort levels in human-robot interaction experiments. We discuss methodological issues of using the device in an exploratory HRI study where subjects were asked to use the device to indicate their subjective comfort level throughout the experiment. The recorded comfort data were time stamped for synchronization and analysis purposes in conjunction with the video footage to help identify certain situations in the HRI trials where subjects felt uncomfortable. In order to provide a proof-of-concept for the suitability of the handheld Comfort Level Device for HRI studies we analyzed the data for seven selected subjects. These examples show that our method helped identifying robot behaviors that subjects felt uncomfortable with. We demonstrate that the device revealed certain uncomfortable states that are visually hidden. Limitations of the device and possible implications for future work conclude the paper.**

*Index Terms – Human-Robot Interaction, Social Robot, Social Interaction, Comfort Level Device.*

## I. INTRODUCTION

In human-inhabited social environments, behaviours or tasks that robots exhibit or perform will result in certain behaviours or responses of humans. Therefore it is essential to understand the relationship between human and robot behaviours in order to create *social robots that humans feel comfortable with*. The issue of human social acceptance has lead to studies that concentrate on the human-centered perspective, where it is essential to include the human in the loop in order to understand what attributes (i.e. behaviour styles, appearance etc.) of robots elicit interactions that are comfortable from the perspective of humans [1,2,3,4]. The research reported in this paper is part of the European project COGNIRON and studies robot companions in a home setting. While such a robot needs to perform and provide assistance for certain useful tasks [5], it should also behave in a socially acceptable manner.

Two main strategies are commonly used for evaluating human-robot interaction from a human subjects' perspective: 1) questionnaires, e.g. used in [5], and 2) analysis of video footage recording the interactions, e.g. [6,7,8]. The latter is more appropriate for scenarios where e.g. verbal inquiry may be impossible (e.g. in the case of non-verbal subjects) [7,8], too intrusive, or might strongly bias the results [9].

For video analysis in our study, a video annotation tool was used to annotate and catalogue specific behaviours of interest from the video footage. Drawback of video analysis is that it is a very time consuming method and that it requires inter-rater reliability tests. Trained video observers are necessary to perform the video analysis. However, there is no guarantee that they will be able to observe all relevant behaviours, let alone subjects' comfort levels which might, if at all, be revealed through language or subtle cues (e.g. facial expressions or utterances indicating discomfort or comfort). Therefore, 'feeling comfortable or uncomfortable' is not necessarily expressed clearly enough that it can be detected from observing video footage. Individual differences in subjects' *expressiveness*, as well as the problem of being able to monitor the subject's face, body movements and utterances continuously during the experiments have encouraged us to pursue an alternative.

In human-computer interaction and robotics, biofeedback sensors measuring physiological variables such as heart beat or skin conductance etc. have been investigated[1]. However, the signal processing required for detecting affect and other internal states is often extensive and sensors need to be attached to the subject. Deriving a high-level concept such as 'comfort' from rich physiological data is not straightforward, although subjects are very familiar with assessing their own subjective 'comfort level'. Thus, we decided to try to directly measure a subject's comfort level via a simple device where subjects use a continuous scale to judge their current comfort level throughout an HRI interaction trial. This led us to the posing of two research questions addressed in the present paper:

RQ1: *Can a simple handheld device be used as a tool for helping researchers identify subjects' comfort level?*

RQ2: *Can a visually hidden uncomfortable state be identified through the use of the Comfort Level Device?*

---

[1] E.g., D. Kulic and E. Croft, "Estimating intent for human-robot interaction", *Proc. IEEE Int. Conf. on Advanced Robotics*, 2003, pp. 810-815; R. Rani, N. Sarkar, C. Smith, L. Kirby, "Anxiety detecting robotic systems – Towards implicit human-robot collaboration", *Robotica*, 22(1): 85-95, 2004.

## II. HUMAN-ROBOT INTERACTION TRIALS

The exploratory study involved single human subjects in a simulated living room scenario. It was carried out at the University of Hertfordshire premises between July and August 2004. This study was conducted using a commercially available, human-scaled, PeopleBot[TM] robot. The main aim of the study was to evaluate, in a task oriented living room scenario, different social behaviour and interaction styles of the PeopleBot[TM] robot from a human-centred perspective. A sample of 28 adult volunteers was recruited from the University of Hertfordshire, balanced for gender, background, and familiarity with technology. All subjects completed consent forms and were not paid for participation.

### A. Experimental Design
Experimental Setup - The Simulated Living Room
The original room measured 8.5 x 4.75m and was partitioned off at one end to form an area that served as a control area for the Wizard-of-Oz [10,11] operators and provided space for the control, network and recording equipment. The room was decorated as a simulated living room.

### B. The Experimental Procedure
The experiment was supervised by an experimenter who introduced and explained the trials to the subject. Each single subject spent about 50 minutes in the simulated living room with only the robot and the experimenter present who interfered as little as possible with the robot trails. The following phases of the experimental procedure are relevant to the present paper.

*Introduction*: A general welcome phase where the robot was introduced to the subject when they entered the simulated living room. An information sheet was given to the subject to read along with a consent form to be signed, then questionnaires were completed. The robot moved around the room whilst the subject completed these initial questionnaires in order to familiarize the subject with the robot.

*Comfort Level Device*: Before subjects proceeded to the main trial, they were given a Comfort Level Device (Fig. 1) and were asked to try it out and operate it a few times (for calibration purposes and in order to provide an opportunity for the subject to get accustomed to the device[2]). Next, they were told to use it throughout the main trial to indicate their comfort level during the trial (see section III). A subset of the data collected in this way during the trials formed the basis of this paper[3].

---

[2] The handheld device might provide an additional potential source of discomfort. We tried to reduce this effect by allowing time for the subject to get used to device. Any potential additional discomfort is likely to be present during the whole trial, and thus less likely to influence the *changes in the levels of comfort/discomfort* which were our primary concern. Focussing on changes in the comfort levels has a second advantage: it makes the data more independent of any 'moods' that a particular subject might be in e.g. on a particular day, assuming that such moods are persistent over a longer period of time. However, these issues merit further investigation.

[3] In terms of the experimental design of our study, we would like to make the following remarks: It would indeed be interesting to see how subjects in a control group, not using the handheld device, would behave. However, the

*Main Trial*: The main trial consisted of two tasks, a Negotiated Space Task and an Assistance Task. The Negotiated Space Task involved the robot moving in the room while the subject went through a pile of books placed on the table, remembering one title at a time, walking over and writing down each title on the whiteboard. The Assistance task involved the subject sitting at the table, copying the book titles from the whiteboard onto a piece of paper and underlining specific letters with a red/highlighter pen. The robot was responsible for bringing the missing red/highlighter pen to the table. The two tasks were chosen as they match two key scenarios studied in the COGNIRON project [12]. At the end of these two task scenarios, the subject completed a robot personality questionnaire. The Main trial was then repeated.

*Final Phase*: The final phase involved the subjects completing several questionnaires.

## III. RESULTS FROM COMFORT LEVEL DEVICE

We built a handheld comfort level monitoring device that would allow subjects to indicate their internal comfort level during the experiment (Fig. 1).



Fig. 1 Photograph of the handheld Comfort Level Device.

The device uses a slider control, located at one edge of the box, to receive users' comfort level feedback. The slider can be moved easily by the subjects using either a thumb or finger to indicate their comfort level. The slider scale was marked on one end of the slider with a happy face, to indicate the subject was comfortable with the robot's behaviour, and a sad face on the other end, to indicate discomfort with the robot's behaviour. The device used a 2.4GHz radio signal data link to

---

primary purpose of our study was to identify whether the handheld device could be used to relate subjects' subjective judgements of comfort/discomfort with observable behaviour. A group of subjects using other, more sophisticated and expensive (e.g. physiological) devices to identify discomfort could serve as a suitable control group. However, those alternative devices were not available to us, and, it is not clear how to easily deduce comfort/discomfort from physiological data. Asking for vocalisations (e.g. "I don't feel comfortable now", or verbal ratings on a scale from one to ten) did not seem appropriate either since it would have interfered with the reading/writing tasks that the subjects were performing. Also, moving a slider with one finger seemed easier to us compared to the effort required in order to pinpoint verbally exact moments of discomfort. Vocalizations would also not be able to provide fine graded quantitative data. Note, our primary aim is to develop a reliable Comfort Level Device for human-robot trials. Thus, a control group involving human-human interaction, instead of human-robot interaction, did not seem suitable either. Our main motivation was to use a simple, very inexpensive device, that can easily be replicated by any talented person with certain engineering skills, and to propose a simple data analysis technique respectively.

send numbers representing the slider position to a PC mounted receiver, which recorded the slider position approximately 10 times per second. The data was time stamped and saved in a file for later synchronisation and analysis in conjunction with the video material. The data downloaded from the handheld subject Comfort Level Device was saved and plotted on a series of charts. However, unexpectedly, the raw data was heavily corrupted by static from the network cameras used to make video recordings of the session (see Fig. 2). We thus developed a method that can digitally clean up this static noise, explained in the next section.

*A. Noise Filtering*

In this section we describe a simple technique for noise reduction in the data[4]. By carefully analysing the raw comfort data, plotted against time (e.g. Fig. 2), we found that it was difficult to distinguish the static noise from the actual comfort data at certain regions of a plot (e.g. the region at time 14:37:41). To overcome this problem, we decided to spread the data points out by plotting the raw comfort level data along the x-axis that was incremented by one data point per step (see Fig. 3). We performed the same plotting method for the subjects' calibration data (see Fig. 4). Next, by comparing the raw comfort level data with the subject's calibration data, we noticed that the characteristics of the static noise were very different from a natural human sliding movement shown in Fig. 4. The raw comfort data contained a lot of random spikes (which were characteristic of static noise) in addition to what appeared to be the subject's actual comfort level profile.

To filter out these random spikes, we decided to use the user calibration data as a reference to determine a threshold value that can be used in our filtering process to prune these random spikes from the raw comfort data. The threshold value was determined by searching through the calibration data to obtain the maximum difference between two data points. The idea was to use the maximum difference between two data points as a threshold value that represents the actual maximum linear velocity the subjects moved the slider under normal conditions. We assumed that only static noise can cause a difference between two points in the raw data exceeding the threshold value.

By using the threshold value, we then scanned through the raw data and replaced the static noise (e.g. $p_i$) with their previous non-static noise data point (e.g. $p_{i-1}$). Note that the threshold value varies with subjects; therefore it was essential to determine each subject's threshold value separately through their calibration data during the filtering process.

Figure 5 illustrates the actual comfort data profile after the filtering process of the raw data (Fig. 2) using a threshold value of 51.

---

[4] It is not our intention to make a contribution to the field of signal processing which has developed far more sophisticated techniques for noise filtering. Instead, we developed a simple technique that turned out to be sufficient for our particular application.
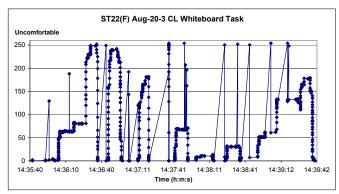


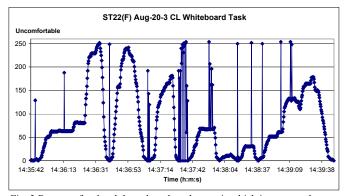Fig. 2 Raw comfort level data plotted against time.



Fig. 3 Raw comfort level data plotted on the x-axis which increments by one data point per step. The time is stamped on the graph every 27 data points.
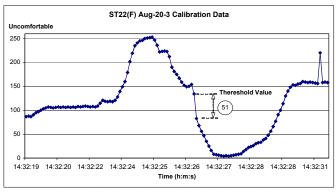


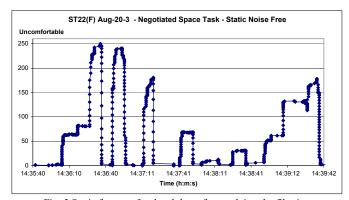Fig. 4 Calibration data indicating the threshold value.



Fig. 5 Static free comfort level data after applying the filtering process using the threshold value shown in Fig. 4.

## B. Analysis of Comfort Level Data

The comfort level data (e.g. Fig. 5) ranged from 0-255, proportional to the motion of the slider, with level 0 representing subjects' most comfortable state (i.e. corresponding to the position of the happy smiley face indicated on the device), while level 255 represents subjects' most uncomfortable states (i.e. sad smiley face).
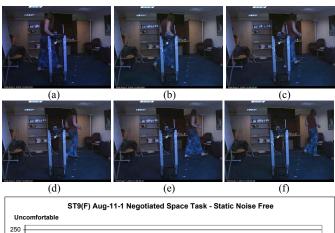
The static free comfort level data of all 28 subjects were visually inspected and classified by the researchers. The data of seven subjects was considered to be very reliable: they clearly used the device consistently and the comfort data ranges from very comfortable to very uncomfortable, so we selected their comfort level data and video data for the present proof-of-concept analysis.

During the initial inspection of the comfort level data (backed by video observation), we found that the majority of the subjects forgot to use their Comfort Level Device after their first interaction task (i.e. after the Negotiated Space Task), see discussion section. For consistency, we decided in this study to concentrate only on the Negotiated Space Task. The fact that only some of the data was suitable for the analysis was not unexpected: a) this was the first time that the newly built device had been used in complex and live HRI trials, and b) this study was our first attempt to gain experience in difficult technical (e.g. interference) as well as methodological issues involved (e.g. how to remind subjects to use the device). We also expected from the outset that the device would only be suitable for particular tasks, we did not expect that the device could be applied generically across the range of all possible HRI scenarios.

For analysing the comfort data, we compared subjects' comfort level data with their corresponding behaviour shown in the experiment (recorded on video). We found that many of the recorded subjects' uncomfortable states corresponded to video sequences where subjects can be either seen moving the slider on the Comfort Level Device, or they were in a difficult situation such as crossing path with the robot, or the robot moving behind them while they were busy writing on the whiteboard. This suggests that a) subjects were willing and able to use the Comfort Level Device, at least in the Negotiated Space Task, and b) the comfort level data had not been produced randomly, but was correlated with subject's behaviour. These correspondences of video data and filtered comfort level data also indicated that the filtering process was successful in filtering out the noise while preserving the subjects' comfort profile recorded during the experiment. This confirms our first Research Question RQ1 - subjects did use the Comfort Level Device to indicate their discomfort. For future trials, it is intended to incorporate error checking and data verification into the RF data transfer link to the recording PC in order to further reduce problems with static.

## IV. VIDEO ANALYSIS

By using the time stamps on the static free comfort data as a reference, we then matched the subjects' uncomfortable states with their video footages recorded during the experiments in order to determine exactly which types of robot behaviours caused the subjects to feel uncomfortable.

Figure 6, a, b, and c illustrate the first half of a video sequence where a subject and the robot crossed paths (the experimental design specifically encouraged such situations which are very common in human inhabited environments - so a robot should be able to deal with it). Here, the subject indicated her discomfort, through the Comfort Level Device, when the robot was heading towards her. The second half of the video sequence (Fig. 6, d, e and f) illustrates a situation where the subject immediately felt comfortable once she had finished crossing the robot's path. The second peak shown in Figure 6g illustrates the recorded subject's comfort level data for the situation shown in fig. 6 (a)-(f).
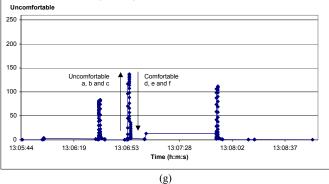


Fig. 6 Video sequences of a human-robot cross path scenario, where the robot stopped and said "after you" as soon as it detected the subject. (a)-(c) illustrate a scenario where the subject indicated that she was uncomfortable with the situation (see g). (d)-(f) illustrate the same scenario where the subject indicated she was comfortable (see g), (g) illustrates subject's comfort level during the Negotiated Space Task.

The comfort level data, along with video footage of all seven subjects revealed that in general there were 3 robot behaviours that were disliked by the majority of the subjects. Firstly, subjects do not like their path being blocked by the robot (Fig. 7a). Secondly, they also found it annoying when the robot moved behind them (Fig. 7b). This situation may be worsened by the robot's sonar sensors which were producing clicking noise that some subjects disliked (as indicated in the final questionnaires). Finally, subjects did not like the robot on a collision path heading toward them in a human-robot cross path scenario (Fig. 7c).

Two out of the seven subjects used only the Comfort Level Device to indicate their discomfort when the robot was

moving behind them (see Fig. 8). These subjects did not exhibit any other physical body language movements to indicate discomfort. This is in contrast to other subjects who used both the Comfort Level Device and body movements such as turning their head to glance at the robot, moving closer to the whiteboard to avoid collision, etc.
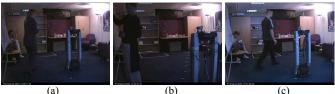
Based on our small sample size we cannot exclude the possibility that the discomfort signals in these situations were produced purely accidentally. However, the striking correspondence with situations where other subjects revealed discomfort, strongly suggests that the Comfort Level Device was used deliberately by the subjects to indicate discomfort. Thus, the Comfort Level Device was able to identify behaviours that are otherwise difficult to be noticed visually (i.e. *visually hidden uncomfortable states*) thus confirming RQ2.

One of the disadvantages of the Comfort Level Device is its sensitivity. We noticed that when the subject (see Fig. 9) opened the whiteboard pen cover, part of his arm motion was transferred to the comfort device slider through his index finger, hence the comfort level data registered 'phantom data' (i.e. registering the subject being in an uncomfortable/comfortable state).

## V. CONCLUSIONS

In this paper we showed that the Comfort Level Device we developed, despite its limitations, was a useful tool that can be applied to the analysis of human-robot interaction, complementing other methods such as video analysis. The simple device turned out to be useful although a) the concept of 'comfort' was not specifically defined, and b) subjects had to 'deliberately' judge their comfort level and reflect their subjective comfort via explicit actions (manual movement of a slider). Before we began the trials it was unclear whether this extra cognitive, as well as manual 'effort' would be accepted by the subjects and yield useful results. However, our results show that the Comfort Level Device provided an insight and feedback from subjects' point of views, revealing which of the robot's behaviours subjects were uncomfortable with.

As expected for a first study using the device, a number of technical as well as methodological problems were identified. The device was suitable for one of the tasks/scenarios studied, but not the other (the majority of the subjects left their Comfort Level Device on the table throughout the Assistance Task). Generally, the device is likely to be more useful for some HRI tasks and contexts than for others. Note, we only reminded the subjects a couple of times during the Negotiated Space Task to use the device. It it thus not surprising that subjects then ignored the device in the second task. Whether the nature of the second task (sitting at a desk and writing) makes it unsuitable for the device, or the lack of reminders to use the device, needs to be investigated further.



(a)      (b)      (c)

Fig. 7 Undesired robot behaviours, a) path blocked, b) robot behind subject, c) collision path.



(a)              (b)

Fig. 8 Visually hidden uncomfortable state, where subjects were feeling uncomfortable but continued writing on the whiteboard. This state was recorded and verified through video observation where subjects were seen moving the slider on the Comfort Level Device.



Fig. 9 Illustration of the phantom effect: uncomfortable data was recorded by the Comfort Level Device when a subject opened a whiteboard pen with both hands while still holding the Comfort Level Device in one hand.

Future work can investigate in more detail the suitability of the device for different scenarios, tasks, user groups[5] etc. However, in this paper we provided proof-of-concept that the device was useful for the data analysis of seven subjects in the Negotiated Space task. Based on our results, it seems that the main issue regarding the Comfort Level Device is not to prove if it is useful of not (we have already shown its usefulness in certain cases), but to map out those HRI scenarios where it can make a significant contribution, in addition to improving on its usability and reliability. Where applicable, the device can replace or complement other devices for measuring subjects' internal states.

Compared to our previous work, which relied solely on observational analysis [7,8], we consider the Comfort Level Device a useful tool. We provided proof-of-concept results for

---

[5] For example, the device is likely not to be suitable for subjects with limitations in manual control or attention.

three selected robot behaviours that the majority of the subjects were uncomfortable with:

a) Robot moving behind subject.

b) Robot blocking subject's path.

c) Robot on collision path with subject.

Subjects' preferred the robot not move behind them, not block their path and avoid being on a collision path (cross path scenario) with them. This situation often occurred when the robot made a turn in the area visibly labelled as 'robot only', leaving the rest of the simulated living room to the subjects. Subjects seemed to prefer the robot not to move around too much when it could interfere with subjects' movements. Also, they did not like to be interrupted in their activities or when the robot got in the subjects' way (i.e. created an obstruction) while subjects were busy with their tasks.

Care should be taken when analysing the Comfort Level Device data to avoid problems such as the phantom data caused by the movements that are caused as a side effect of the subjects' normal body movements and object manipulations.

In terms of our original research questions, we found:

1) A simple handheld device, such as our Comfort Level Device, does provide feedback on subjects' comfort level. We provided proof-of-concept data for seven subjects in the Negotiated Space Task.

2) We identified visually hidden uncomfortable states exhibited by 2 of the subjects which otherwise were very difficult to be identified, even by experienced video observers, without the help of the Comfort Level Device.

Further studies need to confirm the results in this paper using a larger sample size. Currently, we are correlating video data with comfort level data in greater detail in order to support and extend our findings in this paper. Furthermore, we will investigate ways to improve the Comfort Level Device to minimise static noise, reduce phantom data, and find ways to help subjects to continue remembering to use the Comfort Level Device. A very promising direction for future research concerns the possibility that the comfort level data, rather than just being used for post-experimental data analysis and interpretation only, could be used by the robot during the human-robot interaction trials to modify its behaviour style in order to adapt to subjects' preferences, likes and dislikes, an important prerequisite for a personalized robot companion [13].

## ACKNOWLEDGEMENTS

## REFERENCES

[1] T. Saito, T. Shibata, K. Wada and K. Tanie, "Relationship between Interaction with the Mental Commit Robot and 'Change of Stress Reaction of the Elderly", *Proc. IEEE CIRA*, pp. 16-20, 2003.

[2] T. Kanda, H. Ishiguro, "Communication Robots for Elementary Schools", *Proc. AISB'05 Symposium Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction*, pp. 54-63, April 2005.

[3] B. Friedman, P. H. Kahn (Jr.), and J. Hagman, "Hardware Companions? – What online AIBO Discussion Forums Reveal about the Human-Robotic Relationship", *Proc. CHI'03 Conference on Digital Sociability*, vol. 5, pp. 273-279, 2003.

[4] M. Walters, K. Dautenhahn, K. L. Koay, R. te Boekhorst, C. Nehaniv, I. Werry and D. Lee, "Close Encounters: Spatial distances between People and a Robot of Mechanistic Appearance", submitted for publication, July 2005.

[5] K. Dautenhahn, S. Woods, C. Kaouri, M. Walters, K. L. Koay and I. Werry, "What is a Robot Companion – Friend, Assistant or Butler?", to appear in *Proc. IEEE IROS*, Aug 2-6 2005.

[6] T. Salter, R. te Boekhorst, K. Dautenhahn, "Detecting and Analysing Children's Play Styles with Autonomous Mobile Robots: A Case Study Comparing Observational Data with Sensor Readings", *Proc. The 8th Conference on Intelligent Autonomous Systems (IAS-8)*, 10-13 March, Amsterdam, The Netherlands, IOS Press, pp. 61-70.

[7] B. Robins, K. Dautenhahn, R. te Boekhorst, and A. Billard, "Effects of repeated exposure to a humanoid robot on children with autism", *Proc. Universal Access and Assistive Technology (CWUAAT)*, Cambridge, UK, 22-24 March 2004, Springer Verlag (London), pp. 225-236.

[8] K. Dautenhahn and I. Werry, "A Quantitative Technique for Analysing Robot-Human Interactions", *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1132-1138, 2002.

[9] C. D. Kidd, and C. Breazeal, "Human-Robot Interaction Experiments: Lessons learned", *Proc. AISB'05 Symposium Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction*, pp. 141-142, April 2005.

[10] D. Maulsby, S. Greenberg and R Mander, "Prototyping an intelligent agent through Wizard of Oz", *Proc .ACM SIGCHI Conference on Human Factors in Computing Systems*, Amsterdam, The Netherlands, ACM Press, pp. 277-284, 1993.

[11] M. L. Walters, S. Woods, K. L. Koay and K. Dautenhahn, "Practical and Methodological Challenges in Designing and Conducting Human-Robot Interaction Studies", *Proc. AISB'05 Symposium Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction*, pp. 110-120, April 2005.

[12] COGNIRON. Website available at http://www.cogniron.org, 2005.

[13] K. Dautenhahn, "Robots We Like to Live With?! - A Developmental Perspective on a Personalized, Life-Long Robot Companion", *Proc. IEEE Ro-man 2004*, Kurashiki, Japan, IEEE Press, pp. 17-22, 2004.