Empirical Results from Using a Comfort Level Device in Human-Robot Interaction Studies

K.L. Koay, K. Dautenhahn, S.N. Woods and M.L. Walters

University of Hertfordshire

School of Computer Science
College Lane
Hatfield, UK, AL10 9AB
+44 (0) 1707 28 {1026, 4333}

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

ABSTRACT

This paper describes an extensive analysis of the comfort level data of 7 subjects with respect to 12 robot behaviours as part of a human-robot interaction trial. This includes robot action, proximity and motion relative to the subjects. Two researchers coded the video material, identifying visible states of discomfort displayed by subjects in relation to the robot's behaviour. Agreement between the coders varied from moderate to high, except for more ambiguous situations involving robot approach directions. The detected visible states of discomfort were correlated with the situations where the comfort level device (CLD) indicated states of discomfort. Results show that the uncomfortable states identified by both coders, and by either of the coders corresponded with 31% and 64% of the uncomfortable states identified by the subjects' CLD data (N=58), respectively. Conversely there was 72% agreement between subjects' CLD data and the uncomfortable states identified by both coders (N=25). Results show that the majority of the subjects expressed discomfort when the robot blocked their path or was on a collision course towards them, especially when the robot was within 3 meters proximity. Other observations include that the majority of subjects experienced discomfort when the robot was closer than 3m, within the social zone reserved for human-human face to face conversation, while they were performing a task. The advantages and disadvantages of the CLD in comparison to other techniques for assessing subjects' internal states are discussed and future work concludes the paper.

Categories and Subject Descriptors

A.m [**Miscellaneous**]: Human-Robot Interaction – *Social Robots*
I.2.m [**Miscellaneous**]: Robotics – *Mobile Robots*

General Terms

Measurement, Experimentation, Human Factors, Verification.

Keywords

Human-Robot Interaction, Social Robot, Social Interaction, Comfort Level Device.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
HRI'06, March 2–4, 2006, Salt Lake City, Utah, USA.
Copyright 2006 ACM 1-59593-294-1/06/0003...$5.00.

1. INTRODUCTION

Research in the field of Human-Robot Interaction (HRI) has grown over recent years, ranging from human development studies [8][9][15] to therapeutic robots [18][19][20] and robot companions [3][4][6][23][24]. Increasingly robots are able to operate in human-inhabited social environments. It is expected that the behaviour a robot exhibits will elicit certain human responses. Therefore it is not sufficient that future robot's behaviours are intrinsically safe (i.e. Asimov's Three Laws of Robotics [1]) but must also be socially acceptable. This has led to the issues of human social acceptance becoming a growing HRI research area using a human-centered perspective [6][9][19][23] [24].

Our approach is to directly measure a subject's comfort level via a simple handheld device, where subjects use a continuous scale to judge their current comfort level throughout an HRI interaction trial. Two main strategies are commonly used for evaluating human-robot interaction from a human subjects' perspective 1) questionnaires e.g. as used in [4] and 2) video analysis of interactions, e.g. [2][11][18][20][21]. The latter is more appropriate for scenarios where verbal inquiry may not be possible (e.g. in the case of non-verbal subjects) [2][18], too intrusive, or might strongly bias the results [10]. In human-computer interaction and robotics, biofeedback sensors measuring physiological variables such as heart beat or skin conductance etc. have been investigated, e.g. [12][17]. 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'.

In previous work [11] we discussed methodological issues of using a handheld device to assess subjects' subjective comfort levels in human-robot interaction experiments in a simulated living room scenario. We provided proof-of-concept showing that the processed data of the comfort level device can reveal certain situations where the subjects' felt uncomfortable in relation to three selected robot behaviours. We gave examples of situations indicating that the comfort level device can reveal uncomfortable states that were visually hidden, i.e. that cannot be identified by a human observer on video recordings of the interactions. In the present work we did an extensive analysis of the comfort level data of the 7 subjects (whose comfort level data were considered to be very reliable cf. [11]) for the Negotiated Space Task in order to validate the CLD. First, we analysed the comfort level data with respect to 12 robot behaviours, including robot action, proximity and motion relative to subjects. Second, two researchers coded the video material, identifying visible states of discomfort in relation to the robot's behaviour. Agreement between the coders varied from moderate to high except for more ambiguous situations involving robot approach directions. The detected visible states of discomfort were correlated with the situations where the comfort level device (CLD) indicated states of discomfort.

The findings of robot behaviours where subjects experienced discomfort are discussed, followed by the limitations of the CLD, its advantages and disadvantages in comparison to other techniques for assessing subjects' internal states. An outline of future work concludes the paper.

2. HUMAN-ROBOT INTERACTION STUDY

The exploratory study involved single human subjects in a simulated living room scenario. It was carried out at the University of Hertfordshire between July and August 2004. This study was conducted using a commercially available, human-scaled, PeopleBot™ 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™ robot from a human-centred perspective. A sample of 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.

2.1 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 [14][22] operators and provided space for the control, network and recording equipment. Two researchers controlled the robot: one person was driving the robot, the other one was controlling the robot's speech output and the robot's camera movements [22]. Decisions on the robot's movements were based on views from the robot's onboard camera and surrounding cameras located in the room. The room was decorated as a living room in order to provide a comfortable environment relevant to the COGNIRON[1] project which studies a robot companion in domestic scenarios.

2.2 The Handheld Comfort Level Device

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

The device uses a slider control, located at one edge of the box, to receive subjects' 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 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. Figure 2 illustrates a subject using the handheld comfort level device to indicate her comfort/discomfort with the robot behaviours.

Figure 1. Photograph of the handheld Comfort Level Device.

Figure 2. A subject using the handheld Comfort Level Device to indicate her discomfort with the robot's behaviour.

2.3 The Experimental Procedure

The experiment was supervised by an experimenter who introduced and explained the trials to the subject. Each subject spent approximately 50 minutes in the simulated living room with only the robot and the experimenter present who interfered as little as possible with the trials. 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 (Figure 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. The terms comfort/discomfort were not defined prior to the

[1] http://www.cogniron.org

[2] The handheld device might possibly provide an additional potential source of discomfort. By allowing time for the subject to get used to the device we tried to reduce this effect. Any such additional discomfort is likely to persist during the whole trial and is less likely to influence the *changes in the levels of comfort/discomfort* which were our primary concern. Focusing 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. Further studies need to clarify this issue.

experiment, and was left to the subjects' self-interpretation. Therefore the term 'uncomfortable' in the context of this study can be considered as any negative feeling that subjects felt due to a situation they were in that could occur during any intended/unintended interaction with the robot.

A subset of the data collected in this way during the trials forms the basis of this paper.

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. At the end of these two task scenarios, the subject completed questionnaire. The Main Trial was then repeated[3].

Final Phase: The final phase involved the subjects completing several questionnaires. The analysis of the questionnaires and other data collected during the trials is reported in different publications e.g. [4,23].

3. VIDEO ANALYSIS

Video analysis has been used widely in the field of studying human interaction. Its history can be traced back to the early 1940s, where David Efron [5] in his pioneering study of gestures traced his subjects' hands and arms movement through projection of motion-picture films of his subjects frame by frame onto graph paper. The motivation for using video analysis in this study was to identify subjects' **Instances of Discomfort** (IoDs) and associated robot behaviours that caused discomfort during the HRI trials. An annotation tool [13] was used to perform a quantitative evaluation of observational data (video footage of the experimental trials) by first identifying subjects' IoDs, followed by the associated robot behaviours of each IoD. This coding process was performed manually on a second-by-second basis. Subjects' IoDs were identified based on body language, body movements, facial expressions or utterances that indicated discomfort according to the judgment of the coders. This includes e.g. jumpy or jerky body movements, surprised facial expressions and intermittent checking where the robot is.

Details of robot behaviours of interest are discussed in the next subsection.

Table 1. Video Annotation Coding Scheme

Behaviour Code	Action	Robot
A1	Robot Moving behind Subject	Moving behind the subject
A2	Robot Blocking Subject's Path or On Collision Course	• In subject's way or restricting the subject from moving freely • Moving straight at constant speed regardless of subject in its path • Pause forward motion • Change path or break stride across the room
A3	Robot is rotating on the Spot	Rotates in 'robot only area' (a specifically marked area in the experimental room) or area in front of desk
A4	Robot is avoiding subject or avoids getting too close to subject	In the process of avoiding subject (either responding/taking initiative)
A5	Robot is observing subject	Camera pointing ('looking'), tracking or searching for subject
A6	Others	Any observed behaviour that does not fit into any of the behaviour categories defined above
	Proximity	
P1	Close	Less than 1 m
P2	Intermediate	1 to 3 m
P3	Far	More than 3 m
	Motion	
M1	Approach	Moving towards subject's position
M2	Receding	Moving away from subject's position
M3	Stop	Not moving

Note, some of the robot behaviours may occur one after another, but only the first robot behaviour that influenced a particular state of subject's discomfort was coded.

3.1 Video Coding Scheme

The coding scheme for the 12 robot behaviours was produced to help coders identify robot behaviours that subjects were uncomfortable with. This included the robot's actions, as well as proximity and motion relative to the subjects. The 12 behaviours in the coding scheme cover 6 different robot actions, 3 different robot proximities and 3 different robot motions. The 6 different actions and 3 different motions were derived from the robot behaviours identified from the recorded video footage of the experimental trials. The 3 different proximities (distance between subject and robot) used were inspired by Hall's [7] work on

[3] In terms of the experimental design of our study it might be interesting to see how subjects in a control group, not using the handheld device, would behave. However, the main purpose of our study was to identify whether the handheld device could be used to relate subjects' subjective judgments 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, 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 during the experiment. Also, moving a slider with one finger seemed easier to use compared to the effort required 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, using a simple data analysis technique, and validate the approach in HRI experiments.

proxemics[4]. A detailed description of all the 12 robot behaviours is shown in Table 1.

3.2 Video Coding Methods

Two coders were asked to independently code the video annotation using the pre-defined coding scheme to determine the reliability of the coding scheme and identifiable behaviours. They were told to code the robot's behaviour when they noticed the subject exhibiting any signs of discomfort from the video footage. Discomfort in this context refers to the subjects' feeling negative due to any intended or unintended interaction with the robot.

The video annotation process involved the coder going through video footage for 7 subjects who consistently used the CLD, cf. discussion of methodological issues in [11]. The resolution used for annotating the video footage was set at 1 possible IoD per second. For each identified IoD, the coders were required to score the robot's behaviour according to robot action, robot proximity and robot motion relative to the subject, based on the video coding scheme discussed above. Before the actual video coding process began, both video coders were asked to familiarise themselves with the coding scheme using a test video sample.

3.3 Inter-rater Reliability Test (Cohen's kappa)

To determine the consistency of the robot behaviours coded by both independent raters, the two raters' results were compared using a statistical test (Cohen's kappa). From the independent coding of the videos for 7 subjects, coder C_a recorded 52 IoDs and C_b recorded 35 IoDs. Of all the identified IoDs, 25 present in both coders' annotated video data.

These 25 IoDs were used as a basis for the inter-rater reliability tests of agreement for the robot behaviours. The inter-rater agreement between the two independent coders was calculated using the Kappa Coefficient statistics and revealed the following: **Action –** A1, kappa=.364, p=.041 with overall agreement of 72%; A2, kappa=.59, p=.003 with overall agreement of 80%; A6, kappa=.468, p=.006 with overall agreement of 92%; **Proximity** – P1, kappa=.595, p=0.003 with overall agreement of 88%; P2, kappa=.565, p=.004 with overall agreement of 84%; P3, kappa=.648, p=.001 with overall agreement of 96%; **Motion** – M1, kappa=.606, p=.002 with overall agreement of 84%; M2, kappa=.606, p=.002 with overall agreement of 84%.

The majority of the results represent moderate to high agreement with the exception of **Action –** A1 which had a low to moderate inter-rater reliability score. Further analysis revealed that this was mostly caused by ambiguous situations where it was very difficult for the video coder to decide whether the robot's action falls under A1 or A2 (e.g. visually one can see the robot was moving behind the subject, but was the subject's discomfort due to the fact that robot was moving behind her or because the robot was restricting her from moving freely?).

The kappa statistics for **Action** – A3, A5, and **Motion** – M3 were not computed because no such behaviours were recorded by both coders within the 25 matching IoD samples. Kappa statistics was not computed for **Action** – A4 as a symmetric 2-way table in which the values of the first variable match the values of the second variable was not present for this behaviour (within the 25 matching IoD samples).

4. COMFORT LEVEL DATA ANALYSIS

This section discusses the results of the CLD in terms of three different analyses. Firstly, subjects' subjective IoDs (subsection 4.1), and secondly the correlation of subjects' subjective IoDs with their respective IoDs identified by video coders (subsection 4.2). This includes the correlation of subjects' subjective IoDs with their respective IoDs identified by both video coders under the strict-matching rule (subsection 4.2.1), and with either one of the video coders under the relaxed-matching rule (subsection 0.) The correlation results discussed in subsection 4.2 lead to the discussion of the issues of visually hidden IoDs (subsection 4.3) and phantom IoDs (subsection 4.4).

4.1 Subjects' Comfort Level Data

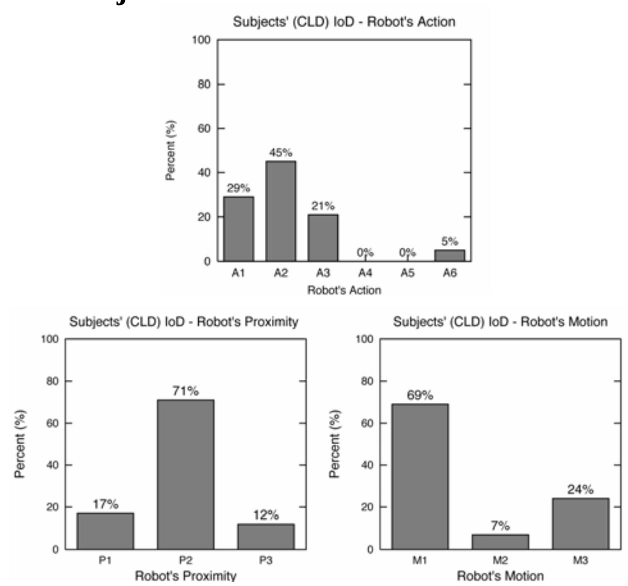

Figure 3. Subject's subjective video annotation results of robot behaviours that subjects were uncomfortable with.

A total of 66 IoDs were obtained from the 7 subjects usage of the CLD. Eight of these were identified as phantom IoDs through the process of video annotation. As discussed in [11] phantom data occurs whenever the CLD's slider is unintentionally moved caused by subjects' body movements, see subsection 4.4. The IoDs in the comfort level data were identified using the techniques described in detail in [11].

The results of the observational data for the robot behaviours based on subjects' comfort level data, excluding the phantom IoDs, revealed that the majority of subjects indicated they were uncomfortable with the robot's action A2 (45%, N=58), the robot's proximity P2 (71%, N=58) around the subjects and robot motion M1 (69%, N=58).

[4] Due to practical reasons (reliability of judging distances from video footage) we could not follow *exactly* Hall's social spaces zones.

4.2 Video Annotation of Subjects' IoDs and Associated Robot Behaviours

In the first phase of the video annotation process, coders C_a and C_b identified IoDs in the video footage of 7 subjects. Findings revealed that the number of IoDs observed by both coders were different: C_a recorded a total of 52 IoDs while C_b recorded a total of 35 IoDs. The breakdown of each coder's observed IoDs for each subject is shown in Table 2.

Table 2. Summary of Subjects' IoD

		Number of Subjcts' IoDs						
		S_a	S_b	S_c	S_d	S_e	S_f	S_g
Video Coder	C_a	6	2	5	15	20	2	2
	C_b	9	1	1	10	10	2	2
Subjects	**CLD**	14	3	3	12	29	3	2
Matching rules	$C_a \cap$CLD	3	2	3	6	11	0	1
	$C_b \cap$CLD	6	1	1	6	8	0	1
	$C_a \cap C_b$	5	0	1	7	9	2	1
	$(C_a \cap C_b) \cap$CLD	3	0	1	5	8	0	1
	$(C_a \cup C_b) \cap$CLD	6	3	3	9	15	0	1
Video Length (min.)		6:32	3:34	3:39	4:29	4:49	4:11	3:14

The difference in the number of visible IoDs identified by both coders was not unexpected, as the detection of subjects' uncomfortable behaviours based on visual cues is a very difficult task, even for trained coders. Factors such as physical restrictions or disabilities may effect the subjects' body movement [16], and can lead to misinterpretation. Also, due to the nature of our unconstrained human-robot interaction experiments, subjects of different heights were allowed to move freely within a reasonably large experimental area to complete a task. Therefore, it was very difficult to have an ideal camera view on the subjects throughout the experiment, despite the use of several video cameras. Note, that the coding of subjects' IoDs may be subject to bias (i.e. coder dependent), even though a pilot test was first introduced before the main video annotation to overcome/minimize the problem of coder dependency. It was also expected that there would be a low percentage of agreement between the matched IoDs for both coders, and the CLD generated IoDs.

Therefore, two matching rules were introduced and applied to the IoDs identified by the coders, and the CLD generated IoDs to assess how they correlated. The two selected rules were 1) Strict-matching Rule – the IoDs identified by both coders agreed with the CLD generated IoDs i.e. $(C_a \cap C_b) \cap$CLD, and 2) Relaxed-matching Rule – the IoDs identified by either of the coders agreed with the CLD generated IoDs i.e. $(C_a \cup C_b) \cap$CLD.

The correlation result from the strict-matching rule was as expected low, scoring just 18 IoDs as opposed to the relaxed-matching rule which scored a total of 37 IoDs, nearly double of the strict-matching rule's score (see Table 2). The robot behaviours that caused these IoDs are discussed in subsection 4.2.1 for the strict-matching rule and subsection 0 for the relaxed-matching rule.

4.2.1 Strict-Matching Rule of Subjects' Comfort Level based on Video Annotation

Overall, 31% (N=58) of the IoDs where subjects indicated they were uncomfortable (during the experiment, through the CLD) matched the video annotation results from both video coders according to the strict matching rule $(C_a \cap C_b) \cap$CLD.

Based on the strict-matching rule between the video observation results annotated by the video coders and the subjects' CLD data, we find 72% (N=18x3) of matching robot behaviours out of all behaviours coded from the video footage. Of these matching behaviours, A2 scored the highest in the robot action category with 67% (N=15), P2 scored the highest in robot proximity category with 92% (N=12) and M1 scored the highest in robot motion category with 100% (N=12). Table 2 shows IoDs for all seven subjects $\mathbf{S_i}$ as generated by the CLD and scored by the two different coders $\mathbf{C_a}$ and $\mathbf{C_b}$.

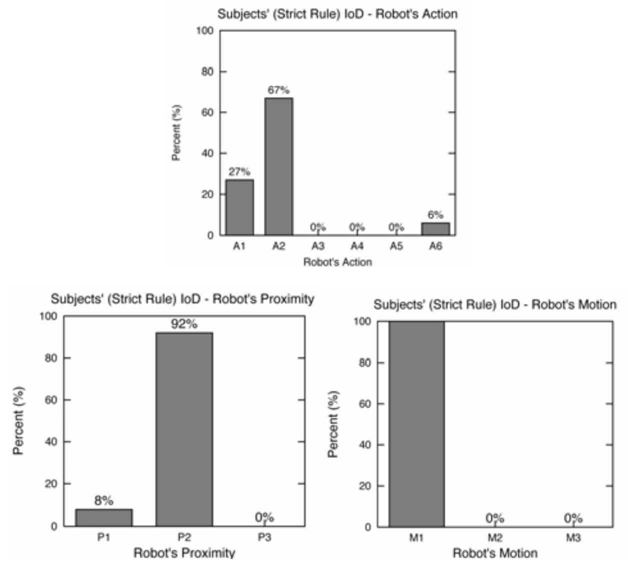

Figure 4. Strict-matching rule video annotation results of robot behaviours that subjects were uncomfortable with.

4.2.2 Relaxed-Matching Rule of Subjects' Comfort Level based on Video Annotation

Overall 64% (N=58) of the IoDs where subjects indicated they were uncomfortable (during the experiment, through the CLD) matched the video annotation results from either of the video coders according to the relaxed-matching rule $(C_a \cup C_b) \cap$ CLD.

Based on the relaxed-matching rule between the video observation results annotated by the video coders and the subjects' CLD data, we find 79% (N=37x3) of matching robot behaviours out of all behaviours coded from the video footage. Of these matching behaviours, A2 scored the highest in the robot action category with 61% (N=31), P2 scored the highest in robot distance category with 69% (N=29) and M1 scored the highest in robot motion category with 82% (N=28).

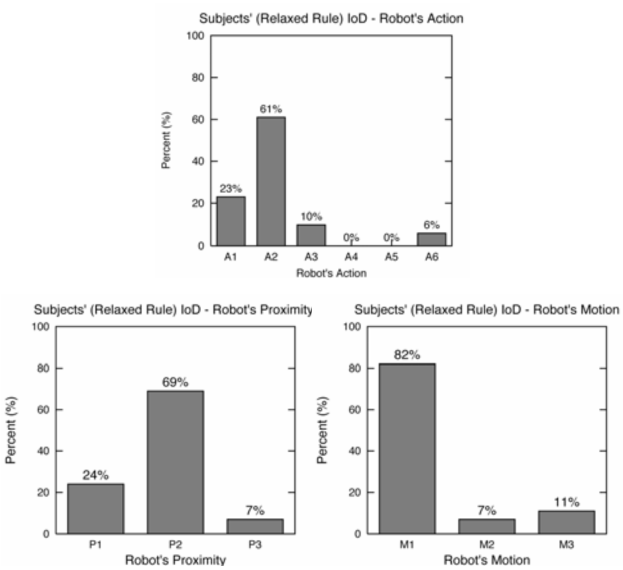

Figure 5. Relaxed-matching rule video annotation results of robot behaviours that subjects were uncomfortable with.

4.3 Visually Hidden IoDs

As shown in Figure 6, a total of 29 visually hidden IoDs were identified from the subjects' CLD data (i.e. IoD's indicated by the subjects via the CLD, but not identified by the video coders). Results from the video analysis on visually hidden IoDs show that 8 of the 29 visually hidden IoDs were actually phantom IoDs (more will be discussed in subsection 4.4). The remaining 21 visually hidden IoDs were observed in four subjects S_a, S_d, S_e and S_f with 8, 3, 7, and 3 IoDs respectively. These visually hidden IoDs suggest three possible explanations: I) Subjects did not always exhibit any visible physical body language movement to indicate their discomfort, II) Subjects did exhibit physical body movements to indicate their discomfort but these were too subtle to be identified by the video coders, and III) due to the nature of the video data in a large unconstrained experimental area, subjects might have been partially or completely out of the cameras' visual field.

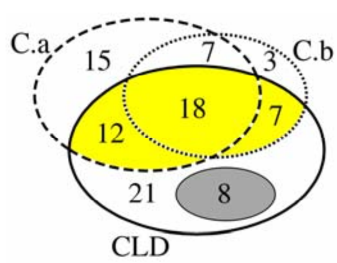

Figure 6. A general overview of the IoDs coded by both coders (i.e. C_a and C_b) and those indicated by the subjects themselves (i.e. CLD). The figure also shows that there were 21 visually hidden IoDs and 8 phantom IoDs.

4.4 Phantom IoDs

Of the 8 phantom IoDs identified, all of them were caused by the subjects' index finger/thumb's unintentional motion resulting from opening and closing the whiteboard pen cover. Seven of the phantom IoDs were identified in the video footage of subject S_e while the remaining 1 IoD was identified in the video footage of subject S_g.

The results from the video observation revealed that the configuration of subjects' grasping/holding the CLD played a pivotal role in the number of generated phantom IoDs (see Figure 7). This was clearly demonstrated for subject S_e, where the majority of the phantom IoDs (i.e. 7 IoDs) came from. Subject S_e was holding the CLD in a configuration where he used his index finger for adjusting the CLD's slider control (i.e. slider control facing the subject's index finger), whereas the other 6 subjects held the CLD in a configuration where they could easily access the CLD's slider control with their thumb (i.e. slider control facing the subjects' thumb) which only generated 1 phantom IoD.

Therefore, it is clear that the problem of phantom IoD we identified in this study could be minimised by advising subjects to grasp the CLD in a particular configuration where the slider control is facing her thumb.

Figure 7. Phantom IoD mainly resulted from the way the subject was holding the CLD.

5. DISCUSSION

The results of the 3 different video analyses discussed in subsection 4.1 and subsection 4.2 revealed 3 different ways of viewing the results of the robot behaviours subjects were uncomfortable with. "Subjects' IoD" which were shown in Figure 3 represents more detailed results of robot behaviours subjects were uncomfortable with based on subjects' CLD data. "Strict-matching rule" which was shown in Figure 4 represents the main robot behaviours subjects were most uncomfortable with when comparing IoD's identified in the video footage and CLD data. "Relaxed-matching rule" which was shown in Figure 5 represents robot behaviours subjects were uncomfortable with including a bigger variety of robot behaviours due to the relaxed-matching rule.

Results show that the uncomfortable states identified by both or either coders corresponded with 31% and 64% of the uncomfortable states identified by the subjects' CLD data (N=58), respectively. Conversely there was 72% agreement between subjects' CLD data and the uncomfortable states identified by both coders (N=25). This was a good result considering coders C_a and C_b coded a total of 52 and 35 IoDs respectively (which was only 6 IoDs less than the subjects' CLD IoDs) for all 7 subjects. This was only a very small subset of all possible 1828 IoDs. This number is equivalent to the total number of time coding steps (that each coder annotated) for all the experiments.

The overall video analysis results of the CLD data from all 3 different analyses (Figure 3, Figure 4, and Figure 5) shows a similar profile and ranking orders of robot behaviours subjects were uncomfortable with. It is difficult to conclude in this paper which robot action subjects were most uncomfortable with, as the live and unconstrained HRI trial does not permit us to design predefined experimental conditions for robot behaviour e.g. with an equal number of robot behaviours in each category. This limits the identification of robot behaviours subjects were most uncomfortable with. Nevertheless, the results give an insight into the dominant robot behaviour from each category, as well as showing which robot behaviours subjects were uncomfortable with in the context of our live and unconstrained HRI trials. As shown in all the 3 different analyses of robot behaviours that subjects were uncomfortable with (see section 4), the dominant robot behaviour for the category Action, Proximity and Motion were A2 (Robot blocking Subject's Path or On Collision Course), P2 (Intermediate proximity – 1 to 3 meters) and M1 (Approach) respectively. The coded robot behaviours (i.e. robot action, proximity and motion) for each associated IoDs were interrelated, but were coded separately into 3 categories. Therefore, the main finding can be interpreted as that a majority of subjects were discomforted when the robot blocked their path or was on collision course towards them; especially when the robot was within 3 m.

The current results confirm previous reported findings [11] (i.e. subjects dislike the robot moving behind them, blocking their path or on collision path toward them.). However, the current study identified that subjects were also uncomfortable with robot action A3 (Robot is rotating on the spot: in an area marked as ‘robot only area’) and A6 (Others: actions that were not defined in the coding scheme which included: 1) when the robot said “excuse me” or “after you” which could interrupt the subject in his task, or the subject did not understand what the robot was saying; 2) the robot was heading towards a possible collision with an obstacle.)

A minority of subjects were uncomfortable even when the robot was 3 meters or more away from them. This could be due to the subject being nervous with the robot moving within their workspace. The majority of subjects experienced discomfort when the robot was closer than 3m. This was understandable, as 3 m is within Hall’s [7] Social Zone, which is mainly reserved for face to face conversation. It was shown in our previous work [24] that human-robot approach distances can be comparable to human-human social distances in some cases. Therefore it is possible that subjects may feel insecure, threatened or intruded upon when a robot enters their personal or social zones.

Results for robot motion indicated that the majority of subjects were uncomfortable when the robot was approaching them. Especially when they were writing on the whiteboard (i.e. robot was moving behind them), or trying to move across the experimental area between the whiteboard and the desk, where the books were located (see Figure 2).

Of the 6 robot actions that were defined in the video coding scheme, only one subject was not comfortable with robot action A3. This may be due to the subject feeling insecure having the robot perform an action behind or out of subject’s visual field (i.e. video observation revealed that the subject was seen occasionally looking at the robot when the robot was performing action A3.) We did not find any indication that the subjects were annoyed by robot’s action A4 (robot is avoiding subject or avoids getting too close to subject) nor A5 (robot is observing subject). This may not come as a surprise as action A4 may be difficult to be identified by the video coders as this action often followed after action A1 or A2. Therefore subjects’ IoDs may have already been associated to either action A1 or A2. Whereas, action A5 was not associated with any of subjects’ IoDs, this may be due to the fact that subjects were concentrating on their tasks and did not notice the robot’s camera, which was quite small in relation to the robot’s physical size, that was observing/tracking them.

In this paper we have shown the usefulness, reliability, and practicality of the CLD, and we believe it is a useful tool to facilitate subjects’ comfort levels in live and unconstrained HRI trials where subjects’ mobility is an essential part of the experiment. We have also shown that the CLD is useful for post-experimental data analysis as the CLD data was able to reveal visually hidden IoDs that were otherwise difficult to be identified by the video coders. The device not only allows the detection of both visible and hidden IoDs, but also answer the problem of video analysis where: (1) subjects might have been partially or completely out of the cameras’ field of view, and (2) subjects are out of focus or far behind in the background of the camera’s field of view (i.e. poor resolution on the focused subjects).

The CLD allows direct identification of subjects IoDs. This is the most crucial and difficult part of video analysis. Even with experienced video coders, there is no guarantee that they will observe all relevant behaviours related to subjects’ IoDs. These IoDs are typically revealed through body language or subtle cues such as facial expressions or utterances.

A limitation of using the CLD is that subjects might forget to use the device. Also, the device may not be useful in experiments where subjects are required to use both their hands to complete their task(s).

Comparing to other biofeedback devices such as galvanic skin response sensors attached to the finger, while they may interfere less with the manual task, stress levels caused by the task and other environmental factors may interfere with the results.

We have highlighted in this paper how the identified phantom IoD problem can be overcome. Like most newly developed technology there is still room for improvement and we expect other unforeseen problems/bugs (e.g. different phantom IoD problem) to surface in our future experiments, which will allow us to improve the CLD.

Future work needs to address the issues of encouraging subjects to use the CLD continuously, not just for extreme behaviours. In future it would be useful if the image of a dimensioned rectangular grid could be superimposed on the experimental floor area in the video recording. This would allow video coders to estimate the robot’s proximity to subjects accurately. As this grid would be purely a video artifact there would be no chance of influencing subjects’ movement during the experiment. A laser scanner might be used instead. More trials need to be carried out with the CLD to identify and overcome any hidden limitations. While in this study we used the CLD for post data analysis, a very promising direction for the future use of the CLD is to allow subjects to modify the attributes of a robot’s behaviours style in order to adapt to subjects’ preferences, likes and dislikes as a requirement for developing a personalized robot companion [3].

6. ACKNOWLEDGEMENTS

The work described in this paper was conducted within the EU Integrated Project COGNIRON ("The Cognitive Robot Companion") and was funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020.

7. REFERENCES

[1] Asimov, I. I, Robot (a collection of short stories originally published between 1940 and 1950), Grafton Books, London, 1968.

[2] Dautenhahn, K., and Werry, I. A Quantitative Technique for Analysing Robot-Human Interactions", Proceedings of IEEE IROS (Lausanne Switzerland, September -October 2002), 1132-1138.

[3] Dautenhahn, K. Robots We Like to Live With?! – A Developmental Perspective on a Personalized, Life-Long Robot Companion. Proceedings of IEEE RO-MAN (Kurashiki Okayama, September 2004), 17-22.

[4] Dautenhahn, K., Woods, S., Kaouri, C., Walters, M., Koay, K.L., and Werry, I. What is a Robot Companion - Friend, Assistant or Butler? Proceedings of IEEE IROS (2005), 1488-1493.

[5] Efron, D. Gesture, Race, and Culture. The Hague: Mouton, 1972. (Originally published as Gesture and Environment. New York: King's Crown Press, 1941.)

[6] Friedman, B., Kahn (Jr.), P.H., and Hagman, J. Hardware Companions? – What online AIBO Discussion Forums Reveal about the Human-Robotic Relationship, Proceedings of CHI'03 (Fort Lauderdale Florida, April 2003), vol. 5, 273-279.

[7] Hall, E.T. The Hidden Dimension: Man's Use of Space in Public and Private. The Bodley Head Ltd, London UK, 1966.

[8] Kanda, T., Hirano, T, Eaton, D., and Ishiguro, H. Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial. Human Computer Interaction (Special issue on human-robot interaction), 19, 1-2 (2004), 61-84.

[9] Kanda, T., and Ishiguro, H. Communication Robots for Elementary Schools. Proceedings of AISB'05 Symposium Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction (Hatfield Hertfordshire, April 2005), 54-63.

[10] Kidd, C.D., and Breazeal, C. Human-Robot Interaction Experiments: Lessons learned. Proceeding of AISB'05 Symposium Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction (Hatfield Hertfordshire, April 2005), 141-142.

[11] Koay, K.L., Walters, M.L., and Dautenhahn, K. Methodological Issues using a Comfort Level Device in Human-Robot Interactions. Proceedings of IEEE RO-MAN 2005 (Nashville TN, August 2005), 359-364.

[12] Kulic, D., and Croft, E. Estimating intent for human-robot interaction. Proceedings of IEEE Int. Conf. on Advanced Robotics (2003), 810-815.

[13] Loehr, D., and Harper, L. Commonplace Tools for Studying Commonplace Interactions: Practitioners' Notes on Entry-Level Video Analysis. Visual Communication, 2, 2 (2003), 225-233.

[14] Maulsby, D., Greenberg, S., and Mander, R. Prototyping an intelligent agent through Wizard of Oz. Proceedings of CHI'93 (Amsterdam, 1993), ACM Press, 277-284.

[15] Michaud, F., Laplante, J-F., Larouche, H., Duquette, A., Caron, S., Letourneau, D., and Masson, P. Autonomous Spherical Mobile Robot for Child-Development Studies. IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, 35, 4 (July 2005), 471-480.

[16] Pease, A. Body Language: How to Read Others' Thoughts by Their Gestures (Overcoming Common Problems). Sheldon Press, 1997.

[17] Rani, R., Sarkar, N., Smith, C., and Kirby, L. Anxiety detecting robotic systems – Towards implicit human-robot collaboration. Robotica, 22, 1 (2004), 85-95.

[18] Robins, B., Dautenhahn, K., te Boekhorst, R., and Billard, A. Effects of repeated exposure to a humanoid robot on children with autism. Proceeding of Universal Access and Assistive Technology (Cambridge UK, March 2004), 225-236.

[19] Saito, T., Shibata, T., Wada, K., and Tanie, K. Relationship between Interaction with the Mental Commit Robot and 'Change of Stress Reaction of the Elderly. Proceeding of IEEE CIRA (Kobe Japan, July 2003), 16-20.

[20] Salter, T., te Boekhorst, R., and Dautenhahn, K. Detecting and Analysing Children's Play Styles with Autonomous Mobile Robots: A Case Study Comparing Observational Data with Sensor Readings. Proceedings of the IAS-8 (Amsterdam The Netherlands, March 2004), 61-70.

[21] Tardif, C., Plumet, M.-H., Beaudichon, J., Waller, D., Bouvard, M., and Leboyer, M. Micro-analysis of social interactions between autistic children and normal adults in semi-structured play situations. Int. J. Behav. Develop, 18, 4 (1995), 727-747.

[22] Walters, M.L., Woods, S., Koay, K.L., and Dautenhahn, K. Practical and Methodological Challenges in Designing and Conducting Human-Robot Interaction Studies. Proceedings of AISB'05 Symposium Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction (Hatfield, April 2005), 110-120.

[23] Walters, M.L., Dautenhahn, K., te Boekhorst, R., Koay, K.L., Kaouri, C, Woods, S., Nehaniv, C., Lee, D., and Werry, I. The Influence of Subjects' Personality Traits on Personal Spatial Zones in a Human-Robot Interaction Experiment. Proceedings of IEEE RO-MAN (Nashville TN, August 2005), 347-352.

[24] Walters, M., Dautenhahn, K., Koay, K.L., Kaouri, K., te Boekhorst, R., Nehaniv, C., Werry, I., and Lee, D. Close Encounters: Spatial distances between People and a Robot of Mechanistic Appearance. Accepted for Humanoids 200