DE GRUYTER

Paladyn, J. Behav. Robot. 2018; 9:95–109

Research Article

Open Access

Kerstin Fischer, Hanna Mareike Weigelin, and Leon Bodenhagen*

Increasing trust in human–robot medical interactions: effects of transparency and adaptability

Abstract: In this paper, we examine trust in a human-robot medical interaction. We focus on the influence of *transparency* and *robot adaptability* on people's trust in a human-robot blood pressure measuring scenario. Our results show that increased transparency, i.e. robot explanations of its own actions designed to make the process and robot behaviors and capabilities accessible to the user, has a consistent effect on people's trust and perceived comfort. In contrast, robot adaptability, i.e., the opportunity to adjust the robot's position according to users' needs, influences users' evaluations of the robot as trustworthy only marginally. Our qualitative analyses indicate that this is due to the fact that transparency and adaptability are complex factors; the investigation of the interactional dynamics shows that users have very specific needs, and that adaptability may have to be paired with responsivity in order to make people feel in control.

Keywords: human-robot interaction, trust, adaptability, medical interaction, blood pressure measurement, transparency

1 Introduction

In a healthcare scenario where the robot takes over routine medical tasks such as heart rate and blood pressure measurements, trust into the system plays an important role. Previous work (e.g. [1]) has identified several factors that influence trust; for instance, the system's performance, such as reliability, false alarm and failure rates [2], plays a role, but also the way the system is introduced [3], and particular functionalities of the system, like its anthropomorphism, transparency, ease of use and politeness [1]. Choi and Li [4, p. 699] suggest that trust depends on three factors:

- system transparency
- technical competence
- situation management

Transparency is a means to make users understand the human-robot interaction situation, the robot's current state and upcoming actions, and the robot's capabilities [5]. Transparency can be communicated in many ways, for instance, via speech, sound, or images, where speech is the most natural form of communication among people (cf. [6]). The way transparency is communicated has so far been addressed in few studies [1], and with rather inconclusive results (see [6]). Technical competence, in contrast, has been addressed in numerous studies (see [2]), and since it is always optimized for, it provides the least grounds for optimization during human-robot interaction. Situation management is mostly concerned with people feeling in control of the situation. In the context of medical measurements, this mostly relates to the ways in which users can adjust the robot to accommodate best to their personal needs, but also the degree to which the robot takes them into account. Thus, of the three factors identified by Choi and Li [4], transparency and adaptability are the two factors that seem to be most promising. In their meta-analysis of the literature on trust in automation, Schaefer et al. [1] explicitly identify communication and transparency as an area where more research is needed. At the same time, their meta analysis indicates that adjustability of the system should increase users' trust [1]. This leads us to focusing on transparency and adaptability as means to increase users' trust into medical robots.

Kerstin Fischer: Department of Design and Communication, University of Southern Denmark, 6400 Sønderborg, Denmark, E-mail: kerstin@sdu.dk
Hanna Mareike Weigelin: University of Southern Denmark, 5230 Odense, Denmark, E-mail: hanna.weigelin@web.de
*Corresponding Author: Leon Bodenhagen: SDU Robotics, Maersk Mc-Kinney Moller Institute, University of Southern Denmark, 5230 Odense, Denmark, E-mail: lebo@mmmi.sdu.dk

2 Previous work

Previous work concerns work on influencing factors of trust in general, and in human-robot medical interactions in particular, as well as work on the two influencing factors investigated, i.e. transparency and adaptability.

2.1 Trust in medical interactions

Hancock et al. [2] conducted a meta-analysis of 29 studies that investigated trust in human-robot interaction. They find that trust development and maintenance were moderately correlated with the performance of the robot. This concerns the reliability, false alarm rate and failure rate of the robot. In contrast, robot attributes like proximity, personality, and anthropomorphism were found to have a minor effect on trust. The authors concluded that robot performance has the strongest influence on users' trust.

Another meta-analysis of 101 papers on trust by Hoff and Bashir [7] distinguishes between three types of trust: dispositional trust, which concerns properties of the respective users, such as age, gender, culture and personality; situational trust, which comprises all context-dependent factors; and learned trust, based on previous experience with the system. Distinguishing different kinds of trust moves the focus from the robot as more or less 'trustworthy' to the interplay between the person trusting, his or her previous experience, the situation, and the respective robot. The authors explicitly mention medical interactions as particularly challenging and suggest that anthropomorphism, transparency, ease of use and politeness should be addressed in future work.

Regarding medical interactions, research on trust in healthcare robots suggests that it mostly depends on three factors: on robot performance expectations, i.e., the perception that the robot will be useful; on social influence, i.e., that others think that one should use a robot, and on the perception that one has sufficient knowledge about the robot and resources to deal with it [8]. Similarly, the use of robot assistance during surgery often depends on the surgeon's own attitude towards the technology in terms of perceived ease of use and complexity, usefulness and behavior control [9]. Large scale surveys further revealed that the potential patient population overwhelmingly welcomes this kind of technology [10, 11].

Experimental research on trust in medical interactions is rare. One reason may be that investigating trust in human-robot interactions experimentally can be ethically challenging since exposing people to situations in which trust matters may mean to either put them in real danger, or to deceive them into believing that they are in real danger [12]. Ximenes et al. [3] address trust not in a medical, but in an invasive scenario with a tattooing robot. In these experiments, the robot uses ink that washes off, but only after some time. The study finds that people trust the robot more when it informed participants before the procedure exactly about what would happen. This suggests that transparency as a factor to influence situational trust could play a major role in increasing trust.

2.2 Transparency

Transparency is usually understood as a unified whole, which can be implemented to different degrees, so that studies on transparency investigate the effects of systems that are more or less transparent (e.g. [13, 14]). More recently, Lyon [5] shows that transparency does in fact comprise different kinds of phenomena. For instance, he argues that in some studies, transparency is understood as users' understanding of why a robot behaved in an unexpected way [13], whereas in others, it is understood as information about the robot's reliability, such as the system's tendency for errors in a given context [15]. Moreover, he argues that the robot also needs to communicate to what extent it is aware of the human and his or her goals. He, therefore, suggests a classification of transparency that comprises both various types of "robot-to-human" transparency, which includes communication about the robot's intentions, task, analytical and environmental model; and "robot-of-human" transparency, which includes communication about the robot's awareness of the human's current states and a model of the kind of teamwork targeted.

Lyons' classification highlights an important point: transparency comprises many different aspects, not all of which may be equally relevant for increasing trust, and some may in fact be counterproductive under some circumstances. For example, in the combat scenario addressed by Lyons, it is absolutely necessary for the robot to communicate when it is not operational or if there is an error of any kind, since human lives rely on a fully functioning robot, then it is indeed crucial to know when the robot may not be reliable. In contrast, for a social robot, which aims to establish some kind of social presence or fulfil a particular social role, like friend, teacher or caregiver, it may be counterproductive to inform the human communication partner when the robot is not working reliably, or even on which basis it may be operating (for instance, using information from a laser range finder of two poles in close proximity to infer that there is a human).

Similarly, in a Wizard-of-Oz scenario [16], in which a robot is secretly controlled by a human operator, the complete truth about the robot's capabilities would not either be desirable. What kinds of information the human user needs to know is thus context-dependent and dependent on the role the robot plays.

To sum up, transparency is not a single variable, but concerns a complex cluster of different kinds of information that may be relevant in different contexts to different degrees.

Moreover, transparency can be reached in different ways, for instance, regarding the kind of delivery of information, previous work has shown that different communication modalities play a role, as a study by Sanders et al. [6] shows; they find that graphically presented information is superior in increasing trust over auditory information, which again is superior over text-based presentations.

But even within the same communication modality, there may be differences in the way in which the information should be communicated. For instance, Fisher [17] compared human-robot interactions with explicit and implicit feedback. She had participants steer a robotic wheelchair to various locations in a room furnished for wheelchair users and label relevant locations, with the aim that the robot should learn to drive the user autonomously to these locations afterwards. Participants were free to decide what would be relevant locations, and to steer the robot to as many locations as they considered necessary. The robot's behavior was scripted and identical across participants and conditions, with one slight difference: in the explicit feedback condition, for three of the locations in the room, the feedback the robot gave was explicit: for instance, when participants steered the robot to the refrigerator and said "this is the refrigerator", in the explicit condition, the robot said "I understood refrigerator - is this where you want to be to open it?," whereas in the implicit condition, the robot only used statements like "Is this where you want to be to open it?". Since participants were free to steer the robot to as many locations as they saw fit, participants heard the explicit robot feedback at most three times in the explicit condition. The effects of these utterances were however considerable: whereas participants showed the robot about ten locations on average in the implicit condition, it were only six in the explicit conditions. Similarly, significant differences were found in the amount of speech directed at the robot, such that participants in the explicit condition interacted much less with the robot. Thus the fact that the robot provided explicit feedback about a low-level capability indicated to the participants that the robot had problems with very basic operations.

For transparency, this may mean that the robot indicating its limitations can lead to unintended inferences and to such low-level mental models of the respective robot that trust is negatively influenced. Furthermore, it can be expected that depending on whether information about the robot is delivered in an implicit or explicit fashion, it will have an impact on how the robot's capabilities are perceived. These two issues render the actual design of transparency non-trivial, and they raise doubts on how honest one may want to be, especially if participants are likely to lose trust into the robot if they understand the real capabilities of the robot.

To sum up, transparency needs to be designed in a context-sensitive way and depending on the role the robot is supposed to play in the specific situation of use. In the kind of medical interactions under consideration, those kinds of information are relevant that inform users about upcoming actions [3] and current states [6]. Since the targeted interaction is social, low level capabilities and limitations of the robot should not be displayed explicitly but rather implicitly.

2.3 Adaptability

Adaptability is suspected to influence users' trust into robots, because they may feel more in control if they can influence the robot's behavior. Choi and Ji [4] find locus of control to be a strong influencing factor on behavioral intention (to use a technology). While transparency only concerns informing participants about the robot, adaptability goes a step further by allowing participants to influence the robot's behavior, which leads to higher trust into the possibility that the human user is still able to control what is going on, especially if the reliability of the system is in doubt. Thus, the ability to control the situation is closely connected to trust into the performance of the system [2].

Ahmad et al. [18] provide a systematic review of adaptivity in human-robot interaction; they analyze empirical studies that address the effects of adaptivity in different domains, such as healthcare and therapy, education, at home and at the work place for the robot capabilities implemented or simulated and the adaptive features used. In the healthcare domain, which is the most relevant for our scenario, they report on nine studies that adjusted their dialog behavior to users' interaction styles, performance, personalities or demographics, or they adjusted the robot's gaming behavior or facial expressions and hand movements to the users' behaviors and emotional displays. Thus, in the studies reviewed, adaptivity concerns personalization and learning more than the op-

portunity to adjust the robot through the user. However, previous work on robot approach shows that if people have the impression that their presence influences the robot's behavior, this puts them at ease (see, for instance, [19], [20] and [21]), i.e., increases their trust. Thus, it can be expected that providing participants with the opportunity to influence the interaction by making the robot adaptable will increase their trust.

2.4 Summary and hypotheses

Lyons criticizes studies in which "novice users are asked to interact with the robotic systems for a brief duration," because in such short interactions, the need for the kind of deep transparency he suggests, does not become apparent. While this is absolutely convincing for close, long-term human-robot collaboration, in brief medical interactions with novice users, trust is equally important. In our scenario, the robot is intended to drive through a care facility and to take over routine medical measurement tasks on a large number of patients and care receivers (see [22]). Thus, the kind of transparency relevant for this scenario concerns that participants understand the robot's next action, its limitations and its current status.

This leads to the following three hypotheses to be tested in this article:

H1 Increasing transparency by explaining current and future actions during the blood pressure measurement procedure has a positive impact on the users' evaluation of the robot. Informing the participant of imminent actions ("I am going to come closer"), of what the user needs to do ("Please tell me when you are done"), and about the status of extended actions ("I am almost finished") is expected to increase users' trust as they know what is happening [5].

H2 Giving the user the option to adjust the position of the blood pressure measurement cuff has a positive impact on the users' evaluation of the robot. Giving participants the opportunity to influence the actions of the robot is expected to make them feel more in control of the situation, which should lead to increased trust [1].

H3 Increasing transparency and allowing adjustment together has an even more positive impact on the users' evaluation of the robot. It is expected that letting the participant influence the robot's actions (as in H2) and providing more clarity of the robot's actions (as in H1) will make the participant trust the robot even more than either of the two factors alone.

Figure 1: The Care-O-bot with a standard blood pressure measurement device attached to its arm.

3 Methodology

The three hypotheses outlined in Section 2.4 are addressed using a between-subject design. The scenario, described in Section 3.1, is a routine medical measurement task in which a robot (see Fig. 1) measures the blood pressure of the participants. That is, the robot welcomes the participants, asks them to sit, drives up to them, asks them to put their arms into the cuff and to close the cuff afterwards, and then inflates the cuff and measures the participants' blood pressure. The study is carried out in four conditions (Section 3.2) to assess the influence of transparency and adaptability. The procedure of the experiment is outlined in Section 3.3 and the participants are described in Section 3.4.

3.1 Experimental set-up

The experiment was conducted in a university lab. The Care-O-bot 3 [23] was chosen for the experiment (see Fig. 1) for the following reasons:

- **Size:** The robot to be used had to be large enough to reach the upper arm of a sitting person;
- **Features:** The robot had to have movable arm to reach the user with the cuff of the blood pressure measuring device;
- **Robustness**: The robot had to be robust enough to withstand any force a user might exert on it during the measurement. This includes not tipping over when the

robot arm is folded out and the cuff is wrapped around the user's arm;
- **Mobility**: The robot had to be able to move across the room to reach the user at a predefined position;
- **Communication**: The robot had to be able to speak in order to communicate with the user efficiently and in a way that is natural to the user;
- **Control**: The robot had to have remote control features so that the Wizard-of-Oz approach is possible.

A conventional blood pressure measuring device[1] was attached to the 6 DoF arm of the robot with velcro fasteners allowing for a quick detachment of the device if needed. The tube between the cuff and the device was extended to allow the operator to start or, if required, abort the measurement without the participant's direct involvement. The robot's behavior was scripted and preprogrammed, yet initiated by a human wizard for safety reasons. The wizard was hidden behind a screen, and participants deemed themselves alone with the robot over the course of the whole experiment.

Participants were recruited among students and staff members by an experimenter and guided to the door for them to enter the room alone, unaware of the hidden operator. The layout of the room is shown in Fig. 2.

3.2 Conditions

The experiment comprises four conditions (summarized in Table 1). In all conditions, the robot offers enough information to enable the participants to complete the experiment. In conditions 2 and 4, however, the transparency conditions, the robot provides further information to increase transparency for the participant. For instance, the robot prepares the participants for the next move by describing its upcoming actions, such as "I am going to come closer now", or "I am going to adjust the height of the cuff based on your body height".

In conditions 3 and 4, the adjustment conditions, the robot ostensibly takes participants into account, for instance, by providing them with the opportunity to have the robot adjust the height of the cuff. Correspondingly, condition 1 constitutes the baseline scenario whereas condition 4 combines the effects of transparency and adjustability. An overview of the conditions is provided in Table 1.

1 Beurer, Oberarm-Blutdruckmessgerät BM 58, https://www.beurer.com/web/de/produkte/blutdruck/blutdruckmessgeraet-oberarm/BM-58
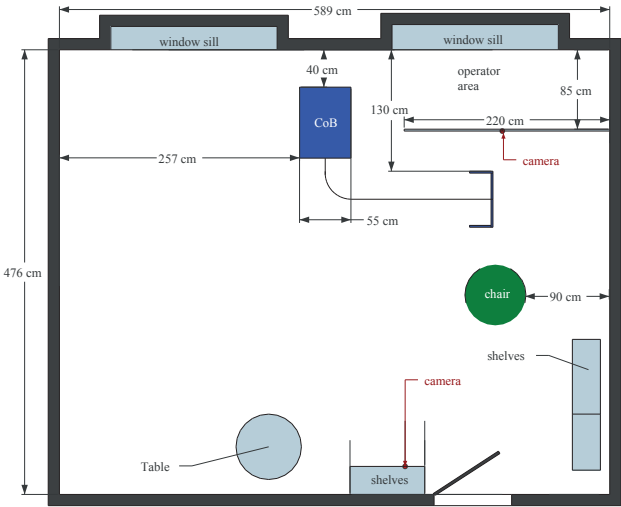

Figure 2: The layout of the room where the experiment was conducted. The participant is seated in the chair (green), the starting position of the robot (CoB) is represented as a blue rectangle, its final position is marked in blue; the operator was hidden behind the screen in the upper right hand corner.

Table 1: The four conditions of the experiment. In all conditions the participants were provided enough information to complete the task.

	No Transparency	**Transparency**
No Adjustment	Condition 1	Condition 2
Adjustment	Condition 3	Condition 4

3.2.1 Transparency

The transparency conditions were designed according to the requirements of the scenario under consideration. In particular, based on previous work, we assume that trust can be increased if the robot explains its upcoming actions and current status so that participants are prepared for what is going to happen. One kind of manipulation thus concerns the announcement of the next action. This is done throughout the experiment:
- before the robot moves to its measuring position next to the participant;
- before folding out its arm;
- before the automatic adjustment of the height of the cuff, based on user height;
- before the potential adjustment of the height of the cuff, based on user input;
- before folding in the arm again.

Furthermore, we do not signal explicitly shortcomings and potential problems of the robot, but rather inform par-

Table 2: Design rationale for the transparency communicated by the robot's utterances

Robot speech	Explanation
"Please tell me when you are ready."	Polite invitation to control the beginning of the interaction (ambiguous between politeness and lack of capability); clear information on what the user is required to do
"I am going to come closer now."	Announcement of future action; simple sentence, indicates that the robot perceives the position of the user
"I am going to move my arm into the standard measuring position."	Preparation for future action; simple, human-centric language; technical terminology to indicate expertise
"I am adjusting the height of the cuff based on your body height."	Announcement of future action; indicates that robot perceives user's height
"Okay, I am lowering / lifting my arm now."	Announcing future action; re-assuring
"I am going to measure your blood pressure now by inflating the cuff."	Announcing future action; simple sentences, yet formal language to indicate expertise
"I have almost finished."	Status description; simple language
"I have determined your blood pressure. I am going to deflate the cuff."	Announcement of status and future action; formal language to indicate professional expertise
"I am going to move my arm back into the travel position."	Announcement of future action; simple language; implicit information about the robot functionality

ticipants when they have to take over, as in "please tell me when you are ready". Such a request would be polite in a human interaction even though the human interactant is perfectly able to see whether the cuff is closed, and whether the participant is ready. The fact that the robot cannot perceive whether the cuff is closed or not, is therefore not explicitly mentioned. Instead, the request to let the robot know when the participant is ready is ambiguous between politeness and inability and thus serves its purpose to inform the participant what they have to do while not drawing attention to the limitations of the robot.

The robot also informs the participant about its current status during an extended action (the blood pressure measurement) as suggested by [6].

Finally, regarding language, we chose simple sentences yet involving technical vocabulary since technical terminology is generally associated with expertise [24]. Furthermore, we use politeness formulas [25]. Table 2 shows the different robot utterances and explains the design rationale for each of them.

3.2.2 Adjustment

In conditions 3 and 4, the robot adapts to the participants' needs. The possibility to adjust the robot is supposed to give the participant a feeling of control over, and influence

on, the actions of the robot. The two manipulations to increase trust employed are first to signal the participant that the robot takes their position into account by adjusting its speed, and second to give participants the opportunity to influence the height of the cuff according to their wishes.

First, concerning approach, the way the robot approaches the participant is adjusted to the presence of the participant. While the robot in the non-adjustment conditions 1 and 2 moves with constant speed, in the adjustment conditions 3 and 4 the robot moves forward, then turns to face the participant, and then continues with a lower speed. Adjustment of robot speed during approach was shown in previous work to put people at ease with an approaching robot since they feel that the robot takes them into account [21]. This is supposed to give the participant the feeling that his or her presence has an influence on the robot. While participants do not have a direct influence on the robot's behavior, adaptation of the approach behavior provides participants with a feeling of being perceived and thus taken into account by the robot.

Second, concerning adjustability, after the user has attached the cuff to his or her arm, the user is asked whether or not the current position is comfortable. If participants are comfortable, the experiment continues, if not, the participant is asked to say whether the arm should be moved up or down. This gives the participant some control over the robot. Since, for reasons of comparability, the adjust-

ment is not supposed to actually make the cuff position considerably more comfortable than it is in conditions 1 and 2, the adjustment is limited to approximately 2 cm. This is enough to be clearly noticeable, yet not enough to make a significant difference in the level of physical comfort.

The robot offers explicitly to adjust the height of the cuff before and after the measurements; other adjustments were not foreseen, such as horizontal adjustments or the abortion of a process once it had begun.

3.3 Procedure

The experiment comprises three parts: a pre-experimental questionnaire, the experiment proper, which is video recorded, and a post-experimental questionnaire.

3.3.1 Pre-experimental questionnaire

The first questionnaire was filled in at the place where the participant was recruited, so that the participants could not hear or see the robot before the experiment. This questionnaire starts out with four questions concerning users' consent regarding the obtained video footage. People who did not agree to be filmed at all could not take part in the experiment and were allowed to leave immediately. The other three questions address the use of the videos in a master's thesis, in publications and at conferences respectively. The following questions elicit demographic background information, including age, gender, height, and disciplinary background, as well as participants' experience with autonomous, self-driving robots and with blood pressure measurement ("How often is your blood pressure measured?"; "Who measures your blood pressure?").

3.3.2 The experiment proper

As soon as the participant entered the room, the operators activated the first sequence of the robot program, and the experiment took place. During the experiment, two cameras were recording video footage so that the operators could activate the different pre-implemented sequences of the robot program at the appropriate time, using the Wizard-of-Oz approach [16, 26] for safety reasons. The actual experiment is divided into four phases (see the appendix for the complete robot behavior description):

- **Entrance:** When the participant enters the room, the robot welcomes the participant, asking him or her to take a seat. Once the participant confirms that he or she is seated, the robot announces its approach (in conditions 2 and 4). In all conditions, the robot approaches the participant and stops next to the participant.

- **Preparation:** In conditions 2 and 4, the robot announces its arm movement. Then, the robot folds out the arm with the cuff and adjusts its height, with or without an announcement, depending on the condition. In all cases, the height of the cuff is adjusted to the height of the participant based on the first questionnaire:
  - Participants shorter than 170 cm receive a low setting;
  - Participants between 170 and 180 cm receive a medium setting;
  - Participants taller than 180 cm receive a high setting.
  After this standard adjustment, the robot asks the participant to attach the cuff to his or her right arm. The participant is asked to tell the robot when he or she is ready. In conditions 3 and 4, the robot then asks whether or not the participant is comfortable, and depending on the participant's answer, the robot adjusts the height of the cuff. As condition 4 combines both transparency and adjustment, this further adjustment is here again announced prior to the movement.

- **Measurement:** The robot announces the start of the measurement, and the operator activates the blood pressure measuring device. In conditions 2 and 4, the robot adds how the measurement takes place and informs the user, once the cuff is fully inflated, that the measurement is almost done.

- **Exit:** After the measurement, in conditions 3 and 4, the robot asks the participant whether or not to adjust the height of the cuff to allow for an easier exit. If the user accepts the offer, the robot may announce the movement (in condition 4) before adjusting the height of the cuff. Regardless of condition, the robot finally asks the participant to remove the cuff and wishes the participant a nice day. He or she then exits the room.

Most experiments took around three and a half minutes from the participant entering until him or her leaving the lab. The actual measurements took around one and a half minutes.

Once the participant had left the room, he or she was handed the second questionnaire.

3.3.3 Post-experimental questionnaire

The second questionnaire consists of 21 questions about the participants' perception of the robot concerning likeability, friendliness, politeness, knowledge, responsibility, sensibility, competence, consciousness, intelligence, trustworthiness, and to what extent it was felt to be reassuring or intimidating. Furthermore, participants were asked how they felt towards the robot in terms of anxiety, agitation, and comfort. While these are subjective ratings without clearcut definitions, we assumed that these concepts together provide an insight into participants' subjective experiences during the interaction. Moreover, participants were asked to what extent they felt they could predict the robot, to what degree they felt that the robot could sense them, to what degree they felt being in control, and to what extent they trusted the robot. For these questions, we used semantic differentials, asking participants to rate their impressions of the robot on a 7-point Likert scale, where 1 corresponds to, for instance, *responsible* and 7 to *irresponsible*. To avoid that the participants answered the questions mechanically, the scales were reversed for every other question. This means that instead of always having all positive attributes on one side, the scales alternate so that participants had to be more alert while answering. The full list of semantic differentials in response to the question "please rate your impression of the robot" are the following:

- Dislike – Like
- Unfriendly – friendly
- Unkind – kind
- Unpleasant – pleasant
- Awful – nice
- Incompetent – competent
- Ignorant – knowledgeable
- Unconscious – conscious
- Irresponsible – responsible
- Reliable – unreliable
- Unintelligent – intelligent
- Foolish – sensible
- Trustworthy – untrustworthy
- Reassuring – intimidating
- Rude – polite

Finally, we asked the participants to what degree they would prefer blood measurement from a robot in comparison with a human caregiver and how likely they would use the robot in the future if it was available to them. The questionnaire closed with an open comment field.

3.4 Participants

85 students and staff from the University of Southern Denmark in Odense participated in the experiment. For several reasons, some participants had to be excluded (see Section 4.1), leaving n=63 participants in total. Of the 63 remaining participants, 49 identified as male and 14 identified as female. Their ages ranged between 20 and 64 years with a mean of 29.5 years (standard deviation 9.7). Participants ranged from 157 cm to 196 cm in height, with a mean of 178.8 cm and a standard deviation of 8.7. All participants were compensated for their time with a bar of chocolate.

4 Results

During the three-week testing period, 85 participants were tested; however, some participants had to be excluded from the analysis, see below.

4.1 Evaluation of the experiments

During the experiments it was found that several participants failed to close the cuff properly, which prevented the cuff from being inflated. Since the robot could not react to this, it just continued as if the cuff had been inflated, we tested whether this failure influenced participants' evaluations of the robot. A total of 16 failed inflations were found, compared to 69 successful inflations. An independent samples T-Test (equal variances not assumed) was performed to determine the effect of the failure to measure blood pressure due to an uninflated cuff. The results can be found in Table 3.

They show that participants who did not experience inflation had significantly less experience with blood pressure measurements. This may indicate that additional explanations about how to apply the cuff would have been necessary for these participants.

Furthermore, as expected, participants who did not experience inflation of the cuff trusted the robot significantly less, found it more unreliable, and more intimidating. They had a significantly lower intention of using the robot again and found it more irresponsible. In addition, near significant differences were detected concerning participants' ability to predict the robot's actions, which was lower without inflation, and their evaluation of the robot's unfriendliness.

Based on these findings, participants who experienced failure concerning the inflation of the cuff were excluded from further analysis. In addition, one participant

Table 3: Statistical results for inflation

	No Inflation	Inflation	
	N = 16	N = 69	
	Mean (SD)	Mean (SD)	t-value p-value
Intimidating	4.56 (1.548)	2.96 (1.344)	t = 3.829 p = 0.001
Trustworthy	4.13 (1.586)	5.51 (1.080)	t = -3.312 p = 0.004
Experience	3.25 (1.483)	4.35 (1.423)	t = -2.688 p = 0.013
Unreliable	3.81 (1.974)	2.43 (1.430)	t = 2.636 p = 0.016
Likelihood of use	3.94 (2.081)	5.28 (1.756)	t = -2.383 p = 0.027
Trust	4.25 (1.770)	5.30 (1.192)	t = -2.266 p = 0.036
Irresponsible	3.38 (1.821)	2.39 (1.427)	t = 2.021 p = 0.057
Prediction	4.19 (2.040)	5.17 (1.414)	t = -1.834 p = 0.083
Unfriendly	2.75 (1.571)	2.01 (1.334)	t = 1.734 p = 0.098

Table 4: Statistical results for restarts w.r.t. the frequency of blood pressure measurements and who usually carries out those measurements, as well as friendliness and politeness

	No Restart	Restart	
	N = 63	N = 5	
	Mean (SD)	Mean (SD)	t-value p-value
BP Frequency	2.06 (0.982)	1.20 (0.447)	t = 3.672 p = 0.007
Unfriendly	2.05 (1.349)	1.20 (0.447)	t = 3.229 p = 0.008
Polite	5.98 (1.486)	6.80 (0.447)	t = -2.978 p = 0.010
BP Execution	2.92 (1.097)	4.20 (0.837)	t = -3.207 p = 0.023

4.2 Analysis of the questionnaire results

To determine whether or not the three hypotheses should be rejected, analyses of the overall effects of the four conditions, of the effects of transparency and adjustment, as well as analyses of age and gender were conducted.

4.2.1 Overall effects

To ensure that a fair distribution of participants had been established, a one-way ANOVA was conducted on age, gender and all other background factors. No significant differences were found for any condition, which suggests that no condition had a disproportionate amount of specific participants.

Afterwards, a one-way ANOVA was performed on the participants' evaluation of the robot in the second questionnaire. Significant results concern perceived comfort $(F(3, 59) = 4.967; p = 0.004)$, and the evaluation of how much participants trusted the robot is near-significant $(F(3, 59) = 2.631; p = 0.058)$.

A *post hoc* Tukey test reveals that both conditions with transparency were perceived as significantly more comfortable than the adjustment condition. In condition 4, in which the robot exhibited both transparency and adjustment, the robot was experienced as slightly more comfortable (M=5.5, SD=1.40) than in condition 1 (M=4.07, SD=1.67)(see Table 5).

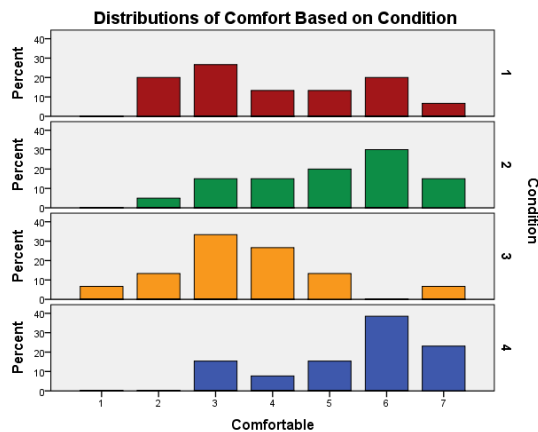Considering trust, the *post hoc* Tukey test shows that people trusted the robot more in condition 4 (M=6.0,

was excluded since he sabotaged the experiment by walking around and not interacting with the robot.

Of the 68 remaining participants, five experienced restarts such that the robot halted its actions and had to be re-initiated by the operator, who therefore sometimes had to give up his cover; in order to determine the effect of restarts, another independent samples T-Test was conducted (see Table 4).

The test reveals that the group of participants who experienced restarts found the robot significantly less unfriendly and more polite. In addition, participants who experienced restarts had their blood pressure measured significantly less frequently and by different people. Because restarts did have an effect on the participants' evaluation of the robot, these participants were excluded from further analysis as well.

This leaves 15 participants for condition 1, 20 participants for condition 2, 15 participants for condition 3, and 13 participants for condition 4, adding up to 63 participants in total.

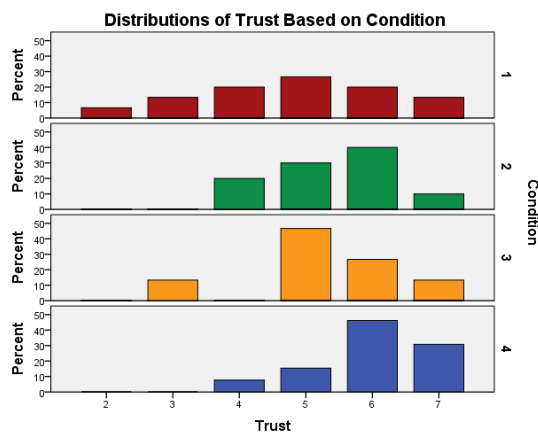**Figure 3:** Distribution of participants' evaluation of how comfortable they were during the experiment

Figure 4: Distribution of participants' evaluation of how much they trusted the robot

SD=0.913) than in condition 1 (M=4.8, SD=1.474) (t = -2.626; p = 0.035). No other factors apart from trust and comfort caused significant or near significant differences between the conditions.

4.2.2 The effects of transparency

To identify the effects of each factor under consideration, a multiple regression analysis was performed for transparency, adjustment, age and gender. The results show that transparency has a significant effect on participants' trust, on their liking of the robot, on the perceived predictability of the robot, on their perceived comfort and on the trustworthiness of the robot (see Table 6). Furthermore, there are near-significant effects for participants' feeling in control and the degree with which they felt the robot could sense them (Table 6). Thus, transparency has

Table 5: p-values from *post hoc* tests for levels of comfort (F = 4.967, df = 3); significant differences are in italics

Comfort	Cond. 1 (N=15)	Cond. 2 (N=20)	Cond. 3 (N=15)	Cond. 4 (N=13)
Cond. 1	-	0.277	0.767	0.080
Cond. 2	0.277	-	*0.030*	0.825
Cond. 3	0.767	*0.030*	-	*0.007*
Cond. 4	0.080	0.825	*0.007*	-

Table 6: Significant and near-significant effects of transparency on the evaluation of the robot

	R^2	F(4,58)	β	p-value
comfort	.211	3.869	.402	*.001*
trust	.118	1.930	.279	*.031*
predictability	.107	1.738	.307	*.018*
trustworthy	.099	1.596	.308	*.018*
liking	.120	1.981	.255	*.047*
sensing	.085	1.344	.227	*.081*
control	.084	1.328	.221	*.090*

a considerable effect on participants' evaluations of the robot.

4.2.3 Effects of adjustment

A regression analysis for the factor adjustment reveals only a near-significant effect on trust (R^2=.1175, F(4,58)= 1.930, β = .225, p= .076); in contrast to transparency, the robot's ability to adjust the height of its arm to participants' wishes did not have a significant effect on the evaluation of the robot.

4.2.4 Gender differences

Previous research shows significant gender differences with respect to the acceptance of technology [27], and thus investigating gender as a potential influencing factor in this study would be relevant; furthermore, Hoff & Bashir [7] find possibly different dispositions of different user groups depending on age and gender to play a major role in trust development.

The regression analysis for gender reveals a significant effect on participants' evaluations such that women felt significantly less relaxed (R^2=.1187, F(4,58)= 1.953, β = -.298, p= .025) and they judged the robot as more sen-
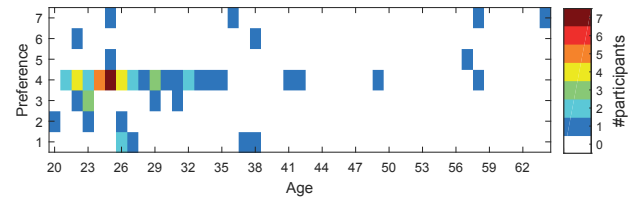
Table 7: Statistical comparison of robot evaluations by gender

	Males N = 49 Mean (SD)	Females N = 14 Mean (SD)	t-value p-value
Irresponsible	2.59 (1.513)	1.71 (0.726)	t = 3.021 p = 0.004
Polite	5.78 (1.571)	6.71 (0.825)	t = -2,983 p = 0.005
Relaxed	5.16 (1.264)	4.14 (1.512)	t = 2.306 p = 0.033
Sensible	5.06 (1.265)	5.86 (1.292)	t = -2.042 p = 0.054

Figure 5: Correlation between age of participants and their preference for human (1) versus robot (7) health assistants

sible (R^2=.071, F(4,58)= 1.116, β = .258, p= .048) and as less irresponsible (R^2=.10524, F(4,58)= 1.710, β = -.264, p= .040); furthermore, there is a tendency for women to feel more agitated (R^2=.1175, F(4,58)= 1.199, β = .239, p= .066) and to evaluate the robot as more polite (R^2=.097, F(4,58)= 1.155, β = .241, p= .061). Thus, women find the robot more polite and less irresponsible, but they are also not as relaxed as men while interacting with the robot (see also Table 7). This is partially in line with the findings of Kuo et al. [27], who suggest that men have a more positive attitude towards robots than women. Nevertheless, concerning most evaluations, there were no significant differences between men and women, which may be partly due to the small sample size of female participants.

4.2.5 Age differences

Concerning potential age-related effects on participants' ratings of trust into the robot, the regression analysis shows a significant effect for participants' preference for having their blood pressure measured by a robot versus by a human care giver (R^2=.124, F(4,58)= 2.059, β = .344, p= .007), as well as a near-significant effect for liking (R^2=.120, F(4,58)= 1.981, β = -.219, p= .083). Figure 5 illustrates the Product Moment correlation between age and preference for human (1) over a robot (7). Quite surprisingly (given the results from large scale surveys, such as the PWC report [11]), several of the older participants in our study indicate that they prefer their blood pressure to be measured by a robot (Age vs. Preference: r(58) = 0.3353; p<.05). In order to exclude that the results are due to the fact that the older participants had more experience with robots, we also calculated the correlations between Age

and Experience (r(58) = 0.1301; p<.05) and between Preference and Experience (r(58) = 0.2250; p<.05). These results indicate that the preference for robotic health assistants by older participants cannot be explained by experience with robots only.

4.3 Qualitative analysis

In the qualitative analysis, the obtained video footage and the comments participants made in the second questionnaire were scrutinized for interaction quality and interaction problems.

4.3.1 Verbal communication

The analysis shows that not everyone was comfortable to interact with the robot via speech. Several participants did not talk to the robot initially but talked to it later, whereas one person never spoke to the robot.

Furthermore, the robot clearly did not speak enough. One person from condition 1 mentioned *"I would maybe have been more calm if the robot have talked more about what he is doing next [...]"(sic)*, which is in line with the quantitative findings showing that transparency is important. The same person also asked that the robot *"[...] maybe talked to me while he was measuring my blood pressure"*. Even in the transparency conditions, the robot only made one comment during the measurement, which participants found to be too little for an interaction of one and a half minutes, as several comments show; participants would have preferred *"assurances that nothing can go wrong"* and *"some small talk like >only 10 more seconds< or >well done<"*.

4.3.2 Robot adjustment

Participants' post-experimental comments reveal that the robot did not adjust enough in the adjustment condition,

which left some participants in an uncomfortable position. Furthermore, the videos show that people did not necessarily want to adjust the height of the cuff in the way the robot offered. For instance, some people wanted adjustment in other directions than the ones offered. One participant changed his mind once he heard that the robot wanted to adjust the height, another one was fine with the adjustment, as long as his arm was not attached to the robot. Other people changed their minds about wanting to adjust, so the robot forced an adjustment onto them as it could not abort a process once it was started.

5 Discussion

The results of our analysis show that transparency, i.e., the robot providing information about its current and upcoming actions, has a considerable effect on people's perceived comfort and trust into the robot. Similarly, transparency positively influenced participants' evaluations of the robot's predictability, trustworthiness, and to some extent their feeling of being in control. Furthermore, participants tended to feel that the robot was sensing them more. Hence, our context-specific transparency interventions were indeed suited to influence positively participants' trust into, and perceived comfort of, the robotic system. Nevertheless, given participants' post-experimental comments and the analysis of the videos, participants would have wished to receive even more of such information. This is in line with the work by Sanders et al. [6], who found continuous feedback to increase trust. Since an interaction with long sequences of silence is very unusual in human interaction (e.g. [25]), the simple fact that the robot was talking at all may have comforted participants, so that not so much transparency, but rather social presence may raise trust into the robot. A future study should thus address what the possible limits are to signalling transparency information during human-robot interaction. In any case, we may conclude that hypothesis H1 can be confirmed.

However, our choice of technical terminology to convey expertise may have not been the best since so many participants failed to close the cuff. Obviously, much more detailed information about what exactly the participant needs to do would have been useful.

Regarding robot adjustability, only a weak effect of adjustability on trust could be found, in contrast to previous findings; this means that hypothesis H2 cannot be confirmed. A possible reason for the failure to match the predictions from the literature may be that Choi and Li [4] used a questionnaire to back up their results, not actual experiments, so that there may be a discrepancy between expected effect and real effect. Regarding other studies that report increased trust due to some level of cooperation (cf. [1]), the adjustments the robot could make in our study may have simply not been sufficient in order to provide participants with a feeling of being in control. Another possibility is that the robot was not responsive enough to give participants the feeling of being in control; comparing the speed and seamlessness of human interaction (e.g. [25]) with the (slow) speed of the current interactions, it is possible that a faster response to participants' requests as well as the ability to abort a process at the moment at which the participant asks for it, may be necessary in order for people to experience the robot as adaptable and for themselves to feel in control (see also [28]). Furthermore, the qualitative analysis has revealed that the robot should have provided a much broader range of options for adjustment than those anticipated, and that inter-individual differences may play a role. Nevertheless, the results also show that together with transparency, adjustability does have a positive effect on trust and a near-significant effect on perceived comfort. Based on these results, hypothesis H3 should not be rejected.

Our analysis of gender differences indicates that while women evaluated the robot more favourably in terms of social traits like politeness, responsibility and sensibility, they also felt more tense around the robot. This is in line with previous research on gender differences in HRI, for instance, [27]. The age differences, in contrast, were relatively unexpected since much work in HRI suggests that younger participants are more willing to adopt new technologies. In any case, most participants did not have a preference for a human over a robot, and eight participants even preferred the robot, which can be counted as a success of the scenario.

Nevertheless, the high number of participants whose blood pressure measurement failed, because they did not close the cuff appropriately, suggests that a much tighter interaction with more dialogue and increased situation awareness would have been desirable. The failure of the robot to measure the participants' blood pressure due to the fact that they had not fastened the cuff properly had a negative effect on trust, perceived reliability and intention of use. Therefore, as in previous work, the robot's (perceived) reliability influenced participants' trust into the robot.

6 Conclusions

To conclude, we found a considerable effect of transparency on users' trust in a human-robot medical measurement scenario, whereas the possibility to make limited adjustments influenced users' evaluations of the robot only marginally. Note, however, that transparency as implemented here was selective, focusing on the signalling of upcoming actions and current activities only, and that we chose for implicit over explicit delivery of information on the robot's capabilities. It is possible that these selections, which were taken to be relevant in the scenario under consideration, are responsible for the positive results and that transparency should be selectively applied. Thus, in the brief interactions in this setting, users profited from information about the robot's next actions, current state and implicit presentation of the robot's capabilities. Nevertheless, the users requested even more verbal robot behavior and re-assurance that everything was going okay.

Transparency and adjustability are hence not simple, binary variables that are either there or not. Instead, they have to match users' specific needs in order to influence users' trust into the robotic system, and both transparency and adjustability will have to be adapted to the situation-specific requirements of the specific context in which the robot is used.

References

[1] K. E. Schaefer, J. Y. C. Chen, J. L. Szalma, P. A. Hancock, A Meta-Analysis of Factors Influencing the Development of Trust in Automation, Human Factors, 2016, 58(3), 377–400, PMID: 27005902

[2] P. A. Hancock, D. R. Billings, K. E Schaefer, J. Y. C. Chen, E. J. De Visser, R. Parasuraman, A meta-analysis of factors affecting trust in human-robot interaction, Human Factors: The Journal of the Human Factors and Ergonomics Society, 2011, 53(5), 517–527

[3] B. H. Ximenes, Í. M. Moreira, J. Kelner. Extreme human-robot interfaces: Increasing trust and assurance around robots, In: The 23rd IEEE International Symposium on Robot and Human Interactive Communication, 2014, 1006–1011

[4] J. K. Choi, Y. G. Ji, Investigating the importance of trust on adopting an autonomous vehicle, International Journal of Human-Computer Interaction, 2015, 31(10), 692–702

[5] J. B. Lyons, Being transparent about transparency: A model for human-robot interaction, AAAI Spring Symposium Series, 2013, 48–53

[6] T. L. Sanders, T. Wixon, K. E. Schafer, J. Y. C. Chen, P. A. Hancock, The influence of modality and transparency on trust in human-robot interaction, IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support, 2014, 156–159

[7] K. A. Hoff, M. Bashir, Trust in Automation: Integrating Empirical Evidence on Factors that Influence Trust, Human Factors: The Journal of the Human Factors and Ergonomics Society, 2015, 57(3), 407–434

[8] A. Alaiad, L. Zhou, G. Koru, An Empirical Study of Home Healthcare Robots Adoption Using the UTUAT Model, Transactions of the International Conference on Health Information Technology Advancement, 2013, 2(1), 185–198

[9] C. BenMassoud, H. Kharrazi, K. F. MacDorman, Facilitators and Barriers to Adopting Robotic-Assisted Surgery: Contextualizing the Unified Theory of Acceptance and Use of Technology, PlosOne, Public Library of Science, 2011, 6(1), 1–11

[10] A. R. Lanfranco, A. E. Castellanos, J. P. Desai, W. C. Meyers, Robotic Surgery: A Current Perspective, Annals of Surgery, 2004, 239, 14–21

[11] PricewaterhouseCoopers, What doctor? Why AI and Robotics will Define New Health, 2017

[12] M. Salem, G. Lakatos, F. Amirabdollahian, K. Dautenhahn, Towards Safe and Trustworthy Social Robots: Ethical Challenges and Practical Issues, International Conference on Social Robotics, Springer, 2015, 584–593

[13] T. Kim, P. Hinds, Who Should I Blame? Effects of Autonomy and Transparency on Attributions in Human-Robot Interactions, 15th IEEE International Symposium on Robot and Human Interactive Communication, 2006, 80–85

[14] T. Chen, D. Campbell, F. Gonzalez, G. Coppin, Management of heterogeneous UAVs through a capability framework of UAV's functional autonomy, 15th Australian International Aerospace Congress, 2013, 25–28

[15] M. T. Dzindolet, S. A. Peterson, R. A. Pomranky, L. G. Pierce, H. P. Beck, The Role of Trust in Automation Reliance, Journal of Human-Computer Studies, 2003, 58, 697–718

[16] L. Riek, Wizard of Oz Studies in HRI: A Systematic Review and New Reporting Guidelines, Journal of Human-Robot Interaction, 2012, 1(1)

[17] K. Fischer, How People Talk with Robots – Designing Dialog to Reduce User Uncertainty, AI Magazine, 2011, 32(4), 31–38

[18] M. I. Ahmad, O. Mubin, J. Orlando, A Systematic Review of Adaptivity in Human-Robot Interaction, Multimodal Technologies and Interaction, 2017, 3(14), 1–25

[19] A. D. Dragan, K. C. T. Lee, S. S. Srinivasa, Legibility and predictability of robot motion, ACM/IEEE International Conference on Human-Robot Interaction, 2011, 301–308

[20] K. Fischer, L. C. Jensen, S-D. Suvei, L. Bodenhagen, Between Legibility and Contact: The Role of Gaze in Robot Approach, IEEE International Symposium on Robot and Human Interactive Communication, 2016, 646-651

[21] M. Lohse, N. van Berkel, E. M. A. G. van Dijk, M. P. Joosse, D. E. Karreman, V. Evers, The influence of approach speed and functional noise on users' perception of a robot, IEEE/RSJ International Conference on Intelligent Robots and Systems, 2013, 1670–1675

[22] K. Fischer, L. Bodenhagen, N. Krüger, M. Ø. Andersen, T. Baumann, Trinvise og delte strategier for kontrol af velfærdsrobotter, Medicoteknik, 2015, 2(1), 11-13

[23] B. Graf, U. Reiser, M. Hägele, K. Mauz, P. Klein, Robotic home assistant Care-O-bot 3 – product vision and innovation platform., IEEE Workshop on Advanced Robotics and its Social Impacts, 2009, 139–144

[24] J. Haard, M. D. Slater, M. Long, Scientese and ambiguous citations in the selling of unproven medical treatments, health communication, 2004, 16(4), 411–426

[25] S. C. Levinson, Pragmatics, Cambridge University Press, 1983

[26] A. Weiss, R. Bernhaupt, D. Schwaiger, M. Altmaninger, R. Buchner, M. Tscheligi, User experience evaluation with a Wizard-of-Oz approach: Technical and methodological considerations, 9th IEEE-RAS International Conference on Humanoid Robots, 2009, 303–308

[27] I. H. Kuo, J. M. Rabindran, E. Broadbent, Y. I. Lee, N. Kerse, R. M. Q. Stafford, B. A. MacDonald, Age and Gender Factors in User Acceptance of Healthcare Robots, The 18th IEEE International Symposium on Robot and Human Interactive Communication, 2009, 214–219

[28] T. Baumann, F. Lindner, Incremental Speech Production for Polite and Natural Personal-Space Intrusion, 7th International Conference on Social Robotics, 2015, 72–82

A Flow charts over the experiment

Fig. 6-8 illustrate the sequence of utterances and actions of the robot during the experiment. Fig. 6 covers the experiment until the measurements starts, Fig. 7 covers the actual measurement, and Fig. 8 covers the end of the experiment including the participant detaching him/herself from the measurement device. Each of the four conditions (see Table 1) is represented by arrows of a specific colour and the colour of the boxes indicate whether it represents a regular utterance/action (green) or if it is tailored for adaptation (yellow) or transparency (blue).
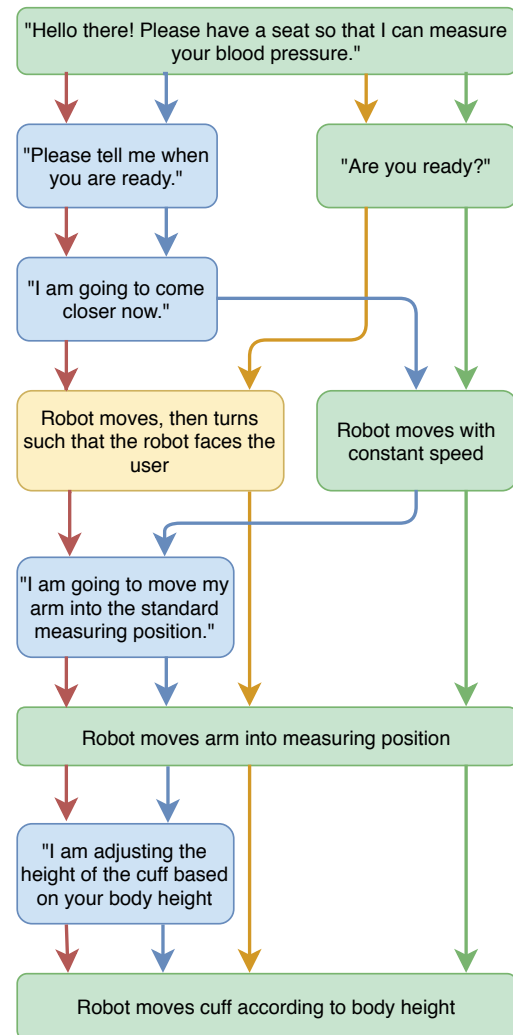

Figure 6: Flow of the experiment, from the start until the measurement. The colour of arrows correspond to the different conditions (green = condition 1, blue = condition 2, orange = condition 3, red = condition 4). The colour of the boxes indicate if the robot behaviour is regular (green) tailored for adaptation (yellow) or transparency (blue).
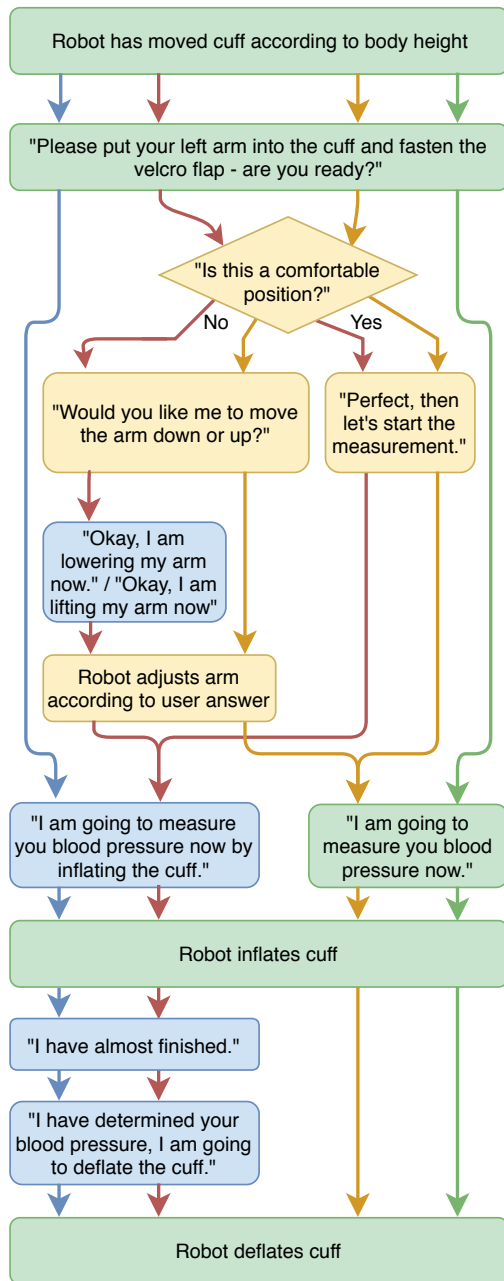
Figure 7: Flow chart of the middle third of the experiment. Behaviour and conditions marked as in Fig. 6.
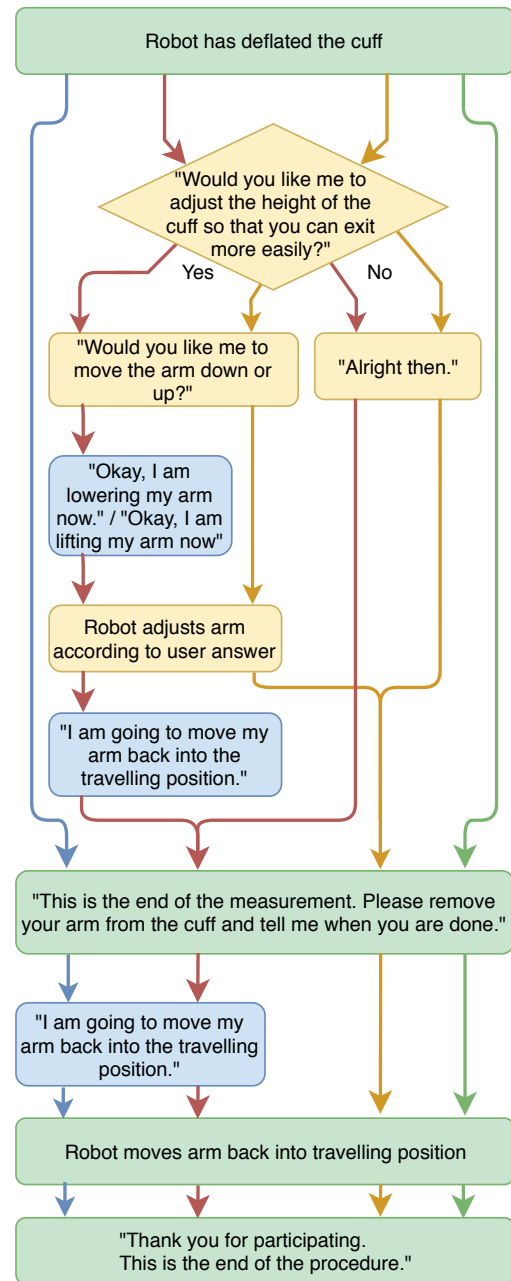
Figure 8: Flow chart of the last third of the experiment. Behaviour and conditions marked as in Fig. 6.