Proceedings Second International Symposium on New Frontiers in Human-Robot Interaction, K. Dautenhahn, J. Saunders (Eds.), at the AISB 2010 convention, 29 March – 1 April 2010, De Montfort University, Leicester, UK

Social, functional, and problem-related tasks in HRI - a comparative analysis of body orientation and gaze

Manja Lohse[1]

Abstract. This paper analyzes human-robot interaction in a home tour scenario with respect to the tasks that the users had to solve in a study. The analysis shows that the users' behavior regarding body orientation and gaze depends on the task which in the present scenario can be social, functional, or problem-related. The paper argues that knowledge about these variations should be taken into account when designing robots in order to improve their capability for situation awareness and, thus, to improve the interaction.

1 INTRODUCTION

The goal to let robots interact with inexperienced users poses grand challenges in human-robot interaction (HRI) today. This is because the users cannot in general be expected to follow concrete plans or to stick to the way of interaction that was foreseen by the developers. However, the theory of situated actions claims that context shapes actions and provides people with means to interpret and understand actions [1]. It is here assumed that tasks in HRI provide one specific kind of context. Therefore, it can be hypothesized that the users' behavior is somewhat similar for specific task whereas it varies between tasks. With this in mind, the current paper investigates the tasks that arise in a home tour scenario. The home tour focuses on teaching rooms and objects to the robot in a domestic environment in order to enable it to solve, for example, fetch-and-carry tasks. In this scenario, the users are likely to be novices or advanced beginners, meaning that they have little or no experience interacting with a personal service robot for the home. In fact, in the user study presented in the following most of the trials were first-contact situations such that the users had never interacted with a robot before.

The main research question in the paper is how the inexperienced users behave during different tasks. To answer this question, the interaction is analyzed applying task analysis (TA) methods. Within the tasks, the paper focuses on the modalities body orientation and gaze. Both these modalities will show whether a differentiation of tasks based on user behavior is actually possible. If it can be found that the users' behavior depends on the tasks, the knowledge about how the behavior varies could be implemented in the robot in order to improve its ability to determine what the users are trying to do. In other words, knowledge about typical user behaviors can be used to improve the robot's situation awareness.

[1] Applied Informatics, Bielefeld University, Universitätsstrasse 25, Bielefeld, Germany. Email: mlohse@techfak.uni-bielefeld.de

2 RELATED WORK

Situation awareness (SA) describes the ability of the robot to find out what is going on in an interaction [2]. In other words, the robot needs to perceive the elements of the environment and to know their meaning. The elements can be analyzed on different levels of specificity of the situation [3]. These levels are (from most general to most specific) situations in general; HRI situations that are characterized by the actors (human and robot); scenarios which determine the roles of the actors (in the home tour the human is the teacher and the robot is the learner); the tasks with specific goals and goal structure; and specific situations (one particular situation). The current paper analyzes elements on the task level by subsuming different specific situations observed in a user study in certain tasks. It looks for similarities within the tasks and differences between the tasks of the home tour. As a preparation for this, the tasks need to be identified with the help of task analysis (TA). TA has been used in market research, product testing, and also in human-computer interaction for a long time. User and task analysis focus on understanding how users perform their tasks [4]. They provide a picture of the system from the perspective of the user, which can be used to improve task completion and to integrate the human into the system design. TA helps to generate hypotheses about sources of inadequate performance of system and operator [5]. In recent years, TA has been introduced to HRI; for example, in the context of situational awareness and user roles [6]. TA has been proposed as a means to identify user's work procedures and tasks when interacting with a mobile office robot, the physical design requirements, function allocation and the relation of the work between user and robot, and user's expectations [7]. The authors approached these goals through interviews and focus groups. In a similar vein, TA has been applied for system design in HRI [8]. None of these authors has focused on TA in the evaluation process, and no data-driven approaches based on user studies have been proposed so far. However, TA is also useful in these respects as the following analyses will show.

Once the tasks are derived, the user behaviors within them need to be analyzed in order to find differences and similarities between the tasks. The user behavior consists of several verbal (speech) and non-verbal (gesture, gaze, spatial behaviors, touch, etc.) modalities which could be used for analysis [9]. Speech is not taken into account here because the users received training on utterances that the robot could understand well. From the non-verbal modalities, body orientation and gaze were chosen because both can actively be produced by the robot and are very important for the interaction.

Spatial behavior in general consists of proximity, orientation, territorial behavior, and movement in a physical setting [9]. All these aspects have been in the focus of the social sciences, often with respect to the influence of gender, age, and culture (for an extensive overview see [10], [11]). In HRI so far especially proximity (spatial distance) was taken into account.

Spatial distances were one main research activity in the EU project Cogniron that focused on "observing a correct attitude towards humans, especially with respect to interaction distances and personal spaces" [12]. In this context, [13] differentiated situations with no interaction, verbal interaction, and physical interaction. In a study they found that the participants wanted the robot to keep the greatest distance when they were not interacting with the robot and allowed it to come closest when interacting physically. The authors argue that in the latter situation the robot's purpose explains the closer approach. In a similar vein, [14] researched situational differences between human preferences for robot approaches in four different scenarios: human seated or standing in the middle of an open space, seated at a table in the middle of an open space, and standing with the back against a wall. All situations are typical for fetch-and-carry tasks. It was found that front left and right approaches were preferred to direct frontal approaches and to rear approaches. Moreover, the participants felt more comfortable with the robot coming closer when they were in open spaces compared to being with their backs against a wall. These results show that contextual and situational factors actually play a role in HRI.

Body orientation also has been of interest for other researchers in HRI. For example, the influence of agent body orientation on task completion was evaluated [15]. However, this work is restricted by the fact that only two orientations were distinguished (facing the user / back towards user). Also Kendon's F-Formation system [16], which is much more natural in interaction, has been discussed [17]. It is based on the observation that people usually group themselves in certain formations when interacting. F-formation describes the kind of formation when people form a shared space between them which they can all access equally well.

Another important modality in social behavior is gaze [9]. The direction of gaze indicates the direction of attention [18]. Thus, where a person looks is crucial. Even though gaze is an important signal for attention, speakers look intermittently at their listeners [9]. One reason for frequent gaze shifts is that looking at the listener all the time could lead to cognitive overload or the arousal caused by the gaze could infer with the cognitive planning. The amount of gaze also varies with the task [9]. For example, tasks like object-teaching require joint attention towards the object [19]. In other words, the interactors look less at each other when there are other things to look at, especially when there is an object of joint attention. Therefore, gaze is important for the comparison of behaviors between tasks.

Some work has also been conducted on gaze in HRI. An experiment was carried out focusing on the production of robot gaze in a scenario where the robot pointed out objects to the participants either correctly or falsely [20]. The measurement of the reaction time of the human depending on the gaze direction and the speech of the robot showed that in the case that the robot's gaze or speech behaviors were inappropriate, the interaction slowed down measurably in response time and fixation distribution. Also others found that non-verbal cues such as gaze convey the robot's internal states which helps to improve human task performance, to prevent and to repair errors [21].

Another study reports how the robot Mel gazed at the participants [22]. In the data, shared looking was coded (mutual gaze and both look at the same object) and a positive relationship between shared looking and engagement of the interactors was found. With respect to the home tour scenario, gaze has been observed to be one mean to establish contact [23]. Users 'look for' feedback in order to find out whether a command has been received correctly by the robot. Therefore, gazing at the robot indicates that some kind of feedback is necessary. It can be concluded that the system should provide continuous information on its status of availability for communication [23]. This status can be provided with the gaze of a camera because just like a human eye the camera behavior of the robot can be used to grab the user's attention and to make contact. Also the robot BIRON exploits this modality as will be described in the following section.

3 ROBOT SYSTEM

In the home tour scenario a domestic robot has to be employed that can be defined "to be a machine that (a) is designed to work with individuals and groups in their personal and public spaces, (b) has a dynamic spatial presence in those spaces, and (c) can "intelligently" interpret its environment and interact physically with it" [24].

BIRON was developed to have these abilities (see Figure 1). The robot is based on an ActiveMedia™ Pioneer PeopleBot platform. It is equipped with a pan-tilt color camera at a height of 141 cm to acquire images of human faces. The camera is also used to show attention by looking in a certain direction and to actively explore the environment visually.

Below the camera, there is a screen which can be used to display the behavioral states of the robot with an animated character called Mindi. The character resembles BIRON. The Mindi pictures are directly related to the robot state, for example, if the robot is in the follow state, an animation is shown in which the Mindi has a happy face and is also walking; if the user is not perceived well, the face of the Mindi character shows an apologetic expression and a speech bubble appears next to the Mindi in which a picture of a fuzzy stick figure is displayed.

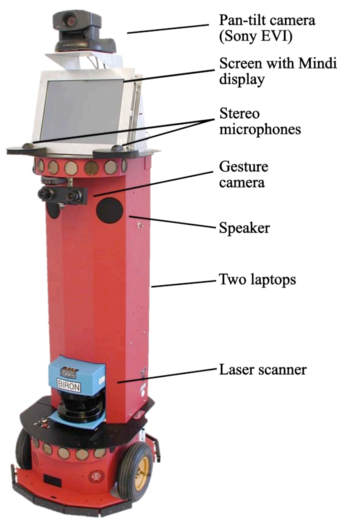

Figure 1. BIRON (BIelefeld RObot companioN)

A pair of AKG far-field microphones is located right below the display at a height of approximately 106 cm. With their help, BIRON can localize speakers. For speech recognition itself, the user wears a headset. Moreover, below the microphones there is an extra camera which is used to recognize gestures, and two speakers for speech output. A SICK laser range finder, mounted at a height of 30 cm, measures distances within a scene.

BIRON is able to follow a person and to autonomously move around. Additionally, it can track people and selectively pay attention to humans looking at it. Further technical information about BIRON is provided in [25].

4 USER STUDY

An exploratory user study was conducted in order to answer the question how different tasks are solved by the users in a home tour scenario. The information gained in the study shall shed light on the question of how the context provided by the tasks influenced the users' behavior with respect to body orientation and gaze.

14 participants (9 male, 5 female; between 16 and 71 years old, mean age 47 years) were invited. Their task was to guide the robot from the living room to the dining room (see Figure 2) and to teach it the book shelf in the living room and the floor lamp in the dining room. All actions that were necessary to conduct this task were practiced with the help of a training sheet beforehand. Thus, training effects were caused on purpose because the participants were completely inexperienced with respect to HRI and needed some basic information in order to decrease insecurity and hesitant behaviors. However, after training no systematic changes of user behaviors were noticed throughout the trials. With the basic information that the participants received, 13 of them completed the whole task. One participant could not finish because of a system breakdown. The completed sessions took between 15 and 33 minutes (average 20 minutes including training).

The whole interaction was videotaped by one of the experimenters. The videos were then used to annotate body orientation and gaze behavior of the users manually. The coding schemes are introduced and validated in [3]. The coding of gaze was very simple, only differentiating three gaze directions (at the robot, at the object in question, somewhere else). In contrast, the coding of body orientation was much more complex. It distinguished static body orientations (neither robot nor human move) and dynamic changes of orientation (either human, robot, or both move). These orientations refer to human and robot. Also the orientation towards the objects of interest was coded based on [26] (see Figure 3). This coding scheme includes one more interval to each side in front of the human than Kendon's F-Formation (codes 1 and 7). Moreover, it is more flexible because it allows to code objects which are behind the human.

The static and dynamic orientations with respect to human and robot used the same intervals as the coding of object positions. However, they also took the facing directions of both interactors into account. While this differentiation might play a minor role in HHI, it is assumed to be important for HRI because of the asymmetry between behaviors that are performed by the human and the robot. Moreover, it was found to be necessary to consider the facing direction of both the human and the robot for two reasons. From the point of view of the robot, the facing direction is crucial because the robot loses the users as soon as it cannot perceive them anymore and the users need to repair this problem before the interaction can continue. Even though the robot is not (yet) programmed to adapt its body orientation to certain tasks based on a social model of the interaction, the teaching of locations requires it to turn around to improve its representation of the environment. From the point of view of the user, the facing direction is important because it indicates the users' expectations about the robot's abilities and their trust in what the system is doing. The users will most probably closely monitor the robot not turning the back towards it if they have doubts about its ability to follow without doing any harm. These considerations resulted in a coding scheme with 23 different codes for static orientations, changes of which were used to code dynamic orientations (see [3]).

The videos were coded with these schemes. If sequences were occluded or unclear they were not taken into account. Reliability of the coding schemes was calculated for two raters to make sure that the video data was suitable to code the modalities and that the coding schemes were usable. The reliability was found to be satisfactory [3].

After having coded body orientation and gaze behavior manually, the annotations were analyzed statistically with the help of SALEM (Statistical AnaLysis of Elan files in Matlab), an approach to do statistical analysis based on annotations in the software Elan ([3], [27]).

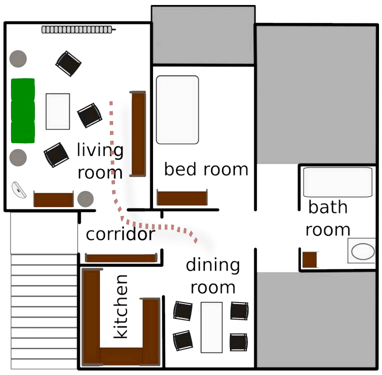

Figure 2. Apartment and path the robot had to be guided

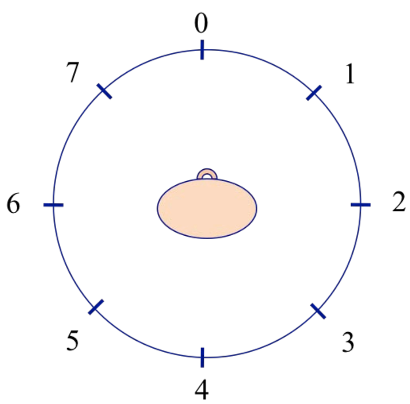

Figure 3. Coding scheme for object position (human in the center; numbers describe object positions)

5 TASK ANALYSIS

In the data of the study, 393 interaction sequences were coded that belong to ten tasks (see Table 1). Each task consisted of certain actions; for example, greeting consisted of the user uttering a greeting which was answered by the robot that also said hello; teaching objects consisted of the user pointing out and naming the object and the robot recognizing the object. Tasks were also coded as such if the user showed a certain intention which the robot did not understand, for example the user greeted the robot which did not reply.

After identifying the single tasks they were grouped according to their function in the interaction. The first group is social tasks. Greeting, intro, and farewell belong to this group. They frame the interaction and are mainly used to manage the robot's attention. On first sight, one could think that greeting and intro could be subsumed into one task. However, they were separated because not every user listened to the intro of the robot, and, more importantly, greetings were also carried out during the interaction to regain the robot's attention while the intro really only occurred after the first greeting. In contrast to the social tasks, guiding, teaching objects and rooms are functional tasks that constitute the main part of the interaction. The functional tasks are mandatory to reach the goal of the interaction. The third group subsumes the problem-related tasks. As the name indicates, these are problems that can take place during the interaction. 'Obstacle' characterizes a situation when the robot cannot move any further because something is in its way. 'Register' describes the state when the robot has lost the person and the user needs to register again by saying hello. The 'reset' task allows the user to reset the robot to a consistent state when it does not react anymore. The 'stop' task is initiated by the robot if the users need to say stop in order to finish a previous task. All these problems, in contrast to mere technical problems, can be solved by the users on the interaction level. The four problem-related tasks could also be seen as part of the other tasks. However, they have been coded separately here for two reasons. Firstly, like the other tasks, they are situations that the user needs to solve using certain behaviors. Secondly, analyzing the problems as tasks allows them to be represented in a transition matrix with all other tasks in order to determine when they occur.

TA revealed that the social tasks took less time than the functional and even as the problem-related tasks. While the introduction of the robot only took place a few times, the greeting task and the farewell task both had very short mean and overall durations. Table 1 also shows that 28.64% of the time was spent on the resolution of problems. The register task alone took up 9.96% of the time of the interaction because it was the most frequent problem. The reset problem did not occur as often; however, it took the longest time to be repaired. This is due to two reasons: firstly, it is the most severe problem and the robot does not tell the user what is wrong and secondly, the robot replies to the reset command that the user has to wait five seconds before saying hello again to reinitiate the interaction. Thus, some time passes without anything happening.

Altogether, most time was spent on the functional tasks (61.26% of the overall duration of the interaction). This is positive because the interaction was designed for these tasks. 34.00% of the time was used for guiding. Guiding is probably the most complex task in the home tour from a technical point of view because it requires the robot to track the the user and the environment while moving. Guiding also caused the most problem-related tasks. The guiding sequences were followed by showing the room, which is the task that should follow it according to the study design, as often as by obstacle and register (transition probability of about .21 each). The obstacle problem is directly related to the guiding of the robot. Obstacle situations mainly resulted when the participants tried to cross the door between the living room and the narrow hallway. That this problem actually interrupted the guiding of the robot can be shown with the high percentage of guiding tasks that succeeded the obstacle task (.88).

Register situations were caused by different problems. The robot lost the users because they moved too fast and too far away from the robot, the users turned too quickly, or the light changed from one room to the other. On the positive side, it can be noted that the problems were hardly ever directly followed by other or the same problems. This shows that the users were actually able to solve them on the interaction level.

Table 1. Descriptive statistics of the home tour tasks

	count	**count of all annotations (%)**	**min. duration (seconds)**	**max. duration (seconds)**	**mean duration (seconds)**	**standard deviation duration**	**duration of all annotations (%)**
social tasks							
greet	45	11.45	0.99	30.23	8.12	6.921	4.99
intro	5	1.27	19.09	57.50	41.81	18.075	2.86
farewell	18	4.58	2.41	23.15	9.13	5.608	2.24
functional tasks							
guiding	109	27.74	3.28	75.07	22.84	12.748	34.00
teaching object	66	16.79	4.14	61.55	19.50	11.583	17.58
teaching room	40	10.18	3.87	37.36	17.71	9.038	9.68
problem-related tasks							
obstacle	25	6.36	8.35	55.38	22.15	9.959	7.56
register	38	9.67	8.19	55.45	19.19	11.548	9.96
reset	18	4.58	13.26	92.14	28.15	19.394	6.92
stop	29	7.38	3.62	29.98	10.60	6.493	4.20
overall	**393**	**100**	**0.99**	**92.14**	**18.63**	**12.719**	**100**

6 ANALYSIS OF BODY ORIENTATION

Now that the tasks have been described, it shall be analyzed how the users behaved while conducting the tasks. For this purpose, body orientation was annotated with the coding scheme that has been introduced in Section 4. As explained there, the static and the dynamic body orientation (i.e., movements that change body orientation) were coded separately. First the results of the static orientation will be introduced.

The analysis showed that the participants turned as often to certain angles at the left side as to the same angles at the right side. Therefore, the orientations which correspond to the same angles, will be grouped for further analysis. In Table 2 it becomes very obvious that the users preferred orientations with smaller angles. They chose them more often and stayed longer in the orientations. The mean durations of the sequences between face-to-face and 45°; and 45° and 90° differed by a factor of about 2.5. Face-to-face orientations were most common which is in line with the findings of [17].

It had further been shown that the body orientation depends on the task that is solved [13]. Here the body orientation was analyzed with respect to all home tour tasks (Table 3). It is obvious at first sight that the users chose a very direct orientation in the social tasks. In more than 80% of the time they maintained a vis-à-vis orientation to the robot. In the remaining time they just turned slightly. This was also true for the problem-related tasks with just a few exceptions. Only in the functional tasks did the participants prefer other orientations toward the robot. Thus, with respect to the coding scheme, it seems that there is in fact a functional difference between the 45° and 90° orientations. While turning to 45° the users can still signal that they are very attentive. However, in the guiding task the probability of turning to a 90° orientation is bigger because the users prepare to walk away from the robot in order to trigger it to follow.

Also the orientation of the users towards the objects was evaluated. Table 4 shows the results of the analysis. As for the coding of the body orientation of the human towards the robot, the amount of times that the users turned to certain angles to their left or right agreed. Whether one or the other was chosen was determined by the fact on which side the object was located. More objects were located to the right of the users, which was true for both the shelf and the floor lamp that most subjects showed to the robot. The angles that the users turned to in order to show these two objects were very similar. Table 4 illustrates that the object was most commonly to the side of the participant's upper body. The object was nearly as often slightly in the back of the user as it was slightly in front of them. This finding indicates that the users did not necessarily navigate the robot until it was in a position where the object was between them as assumed by Kendon's F-Formation system. Figure 4 shows the most common constellations of human, robot, and object.

Table 2. Descriptive statistics of body orientations (overlap (%) = percentage of time the value was annotated in the interaction)

body orientation	count	mean duration (seconds)	SD duration	overlap (%)
face-to-face	169	31.14	43.23	73.71
45°	128	12.21	18.67	21.89
90°	48	5.02	7.26	3.37
135°	12	2.76	1.97	0.46
180°	12	1.59	0.97	0.27
almost side-by-side	1	5.73	0	0.08
side-by-side	1	14.83	0	0.21
overall	**371**	**19.24**	**33.21**	**100**

Table 3. Body orientation in the tasks (# = count, OL (%) = overlap, percentage of time the value was annotated during the task)

	body orientation															
	face-to-face		**45°**		**90°**		**135°**		**180°**		**alm. side-by-side**		**side-by-side**		**overall**	
task	#	OL (%)	#	OL (%)	#	OL (%)	#	OL (%)	#	OL (%)	#	OL (%)	#	OL (%)	#	OL (%)
social tasks																
greet	36	87.55	8	8.26											44	95.81
intro	6	99.11	1	0.34											7	99.45
farewell	12	80.21	5	16.64											17	96.85
functional tasks																
guiding	107	57.89	66	13.25	35	3.22	6	0.63	9	0.56	1	0.19	1	0.52	225	76.25
object	36	46.57	30	39.55	7	7.42	2	0.53							75	94.07
room	35	72.43	11	7.16	3	0.52	1	0.16	2	0.59					52	80.86
problem-related tasks																
obstacle	32	60.87	7	3.55	1	0.44									40	64.86
register	26	60.37	21	25.49	5	2.48	1	0.36			1	0.13	1	0.26	55	89.09
reset	21	63.40	14	29.04											35	92.44
stop	16	77.71	7	14.16			1	1.40	1	0.36					25	93.63

Table 4. Orientations of the users towards the objects

orientation of the object	**count**
between user and robot (0)	1
slightly in front of the user (1, 7)	15
to the side of the user (2, 6)	52
slightly behind the user (3, 5)	13
behind the user (4)	0
sum	**81**

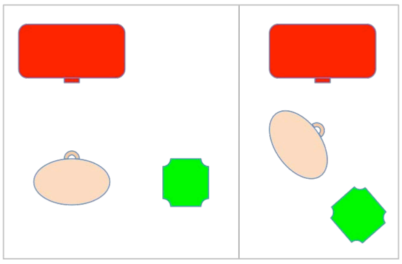

Figure 4. Typical participation frameworks of user, robot, and object in the home tour

Next to the static orientations also changes between different orientations initiated by the users were analyzed. Altogether, 262 movements with a duration of 9 minutes, 36 seconds were coded. The mean duration of the position changes was much shorter than the mean duration of the time that the users spent in a certain position (2.20 seconds vs. 19.11 seconds).

This is underlined by the finding that the users almost always chose the shortest ways between the orientations. But not only the users changed their orientation; the robot also initiated 96 movements that led to a change. This number of changes was significantly lower over all tasks than the number of movements of the users (T-test (two-tailed), df=10, T=2.464, p=.036*). The robot's movements took 7 minutes, 23 seconds altogether. With a mean duration of 4.62 seconds, the movements of the robot were much slower than the movements of the human. The question now is how changes of body orientation relate to the home tour tasks.

The analysis showed that the users changed their orientation during the guiding task most often by far (3.66 movements/ minute). In contrast, the robot most often changed its position in the room-teaching task (1.78 movements/minute), which is in line with the designed robot behavior that makes the robot turn around in order to improve its representation of the room. The task with the second most changes of user (1.97 movements/ minute) and robot (1.48 movements/minute) was the register task. Hence, the users tried to position themselves in front of the robot during the task so that it could perceive them well and the robot turned to the user for the same reason. On the other hand, instead of positioning themselves directly in front of the robot, some users walked to the side of it to see whether it followed, which it did as the number of orientation changes of the robot shows. The users did the same during the reset task (2.01 movements/minute). The reset task was different compared to the register task because the robot did not change its orientation towards the user.

Comparing the movements per minute, the social tasks were found to be rather static (greeting 0.66, introduction 0.57, and farewell 1.10 changes of orientation/minute). The farewell task was accompanied by slightly more movements which is probably due to the fact that the users started to turn away before the robot had replied to the farewell command.

To conclude, the tasks could be differentiated based on the orientation of the users. The users most often changed their orientation in the guiding task which was also due to the fact that the robot was moving. Only during guiding it was at times found that the users turned their backs to the robot. But as has been mentioned above, the time they spent in this orientation was very short compared to the time they spent in more direct orientations. The object-teaching task was the only task during which the participants turned as many times to 45° and 90° (both summed up) as to a vis-à-vis orientation and also spent the same amount of time in these orientations. The finding underlines that the participants turned when teaching objects to establish a participation framework that enables the robot to optimally perceive the objects which are usually located on the side of the participants. This was not found for teaching rooms where the users most of the time faced the robot. Therefore, the actions of teaching rooms and objects seem to follow distinct rules and need to be differentiated.

Altogether, the static and dynamic body orientation have been shown to differ between the tasks and can, therefore, be used to distinguish tasks from each other and to improve the SA of the robot. The following analysis will show if gaze further contributes to the distinction of the tasks.

7 ANALYSIS OF GAZE

Altogether, 2 hours, 13 minutes, and 20 seconds of gaze were analyzed. In 90.7% of the time (491 times) the participants looked at the robot, in 2.7% (159 times) at the objects, and in 6.5% (264 times) somewhere else (see Table 5).

The overall looking time at the robot and the mean duration of glances (14.78 seconds) were much longer than in HHI [9]. Even though the mean duration was strongly influenced by some very long glances, the median was also very high (8.02 seconds). The durations of glances at the objects and at places other than object and robots were much shorter (mean 1.37 and 1.98 seconds respectively, median 1.05 and 1.35 seconds).

The question now is why the users looked at the robot for such a long time. One probable explanation is that they were waiting for feedback which occasionally took long in the study because the robot was acting autonomously. Actually, the longest glances were found in situations when the users were

Table 5. Descriptive statistics of gaze direction

gaze direction	**minimum duration (seconds)**	**maximum duration (seconds)**	**mean duration (seconds)**	**standard deviation duration**	**median duration (seconds)**	**duration (%)**
robot	0.11	259.53	14.78	21.58	8.02	90.7
object	0.38	20.72	1.37	1.76	1.05	2.7
else	0.18	17.74	1.98	1.87	1.35	6.5
overall	**0.11**	**259.53**	**8.69**	**17.08**	**2.47**	**100**

waiting for the robot to reply. This behavior is in line with the finding that gaze at the robot signals that the users need some feedback [23]. Moreover, after the interaction many participants stated that they looked at the screen when the robot did not respond. Since a screen is something one is supposed to look at and not have to look away from once in a while in order to not offend the interaction partner, this is the most probable explanation for the long glances. Unfortunately, from the video data that was recorded, it cannot be determined whether the participants actually looked at the screen or at some other part of the robot. An analysis of gazing behavior during the tasks was conducted to shed some more light on this question (see Table 6).

The first observation in Table 6 is that the number of glances and the percentage of looking at the objects were by far highest in the object-teaching task. This finding is in line with the design of the study and the task. The users most often looked somewhere else other than at the robot or the objects during the intro and the guiding task. During the guiding task the participants certainly looked where they wanted to go to. During the intro the frequent glances away from the robot are more surprising because the robot talks about its usage and often refers to pictures on its screen. When having a second look at the data, it was found that in fact only one user looked away from the robot six times. She turned to the side because the robot had called her by a wrong name (‘Igor’) and she kept laughing about the mistake. Hence, the common gazing direction during the intro really seems to be the robot. But this case also shows that the users might change their behavior crucially if some irregularities appear in the interaction. With respect to the percentage of the time that the users in general looked at the robot, no significant differences were found. Therefore, it can be concluded that gaze direction alone which is here coded in a very simple fashion is only partially revealing. Hence, the analysis needs to integrate several modalities in order to describe user behaviors that are typical for certain tasks. In this context, in the following coherences between gaze direction and body orientation will be described.

The gaze direction corresponded with certain body orientations in the tasks. In the social tasks and the problem-related tasks the body orientation was very direct and the users looked at the robot most of the time. In the guiding task the body orientation as well as the gazing direction changed most often. The users spent most of the time looking somewhere else other than the robot and the object and keeping indirect orientations. The user behavior differed considerably between the two teaching tasks. During object-teaching the users often slightly turned away from the robot and then back towards it as they gazed at the object and then back at the robot. In contrast, during the room-teaching task they gazed at the robot and kept a direct orientation for most of the time. Accordingly, the robot could identify that something was taught based on the utterance and, furthermore, receive information about what was shown via the gaze direction and the body orientation. This finding points out that in the long run even more modalities such as speech could be included in the description of the user behavior during certain tasks.

Table 6. Gaze direction in the tasks (overlap (%) = percentage of time the value was annotated during the task)

	gaze direction							
	at robot		**at object**		**else**		**overall**	
task	count	overlap (%)	count	overlap (%)	count	overlap (%)	count	overlap (%)
social tasks								
greet	50	94.91	1	1.38	7	2.14	58	98.43
intro	14	90.86	0	0	9	9.13	23	100
farewell	20	94.81	0	0	4	4.48	24	99.29
functional tasks								
guide	211	82.07	19	0.90	105	8.28	335	91.25
object	153	87.81	114	**9.85**	18	2.34	285	100
room	60	92.40	12	1.10	26	5.27	98	98.77
problem-related tasks								
obstacle	42	64.19	0	0	8	1.78	50	65.96
register	58	91.19	2	0.17	27	6.41	87	97.77
reset	33	93.03	1	0.08	16	5.45	50	98.56
stop	30	86.41	4	4.87	7	5.71	41	96.99
overall	**459**	**85.75**	**142**	**2.44**	**222**	**5.59**	**823**	**93.79**

8 CONCLUSION

In the previous sections, TA was applied to the home tour tasks in order to evaluate HRI. It was found that the home tour includes social tasks (greet, intro, farewell), functional tasks (teaching rooms and objects, guiding the robot), and problem-related tasks (obstacle, register, reset, stop). These tasks were hypothesized to serve as a kind of context which shapes the users' actions. In order to show that the users' behaviors indeed differed between the tasks, data were analyzed with respect to the modalities eye gaze and body orientation. Coding schemes were introduced for these modalities.

The analysis of body orientation clearly showed that the behaviors indeed differed between the tasks, which allows to distinguish them from each other, for example, a lot of movement and indirect orientations indicate a guiding task while a direct orientation is more common in social and problem-related tasks. With respect to gaze, the analysis showed that the users most often looked at the robot in the social and the problem-related tasks. However, also during the other tasks they mainly looked the robot. Thus, gaze alone was only partially revealing about what the users intended to do. Therefore, it was concluded that the modalities need to be integrated in order to improve the prediction of the users' intention and, in a next step, to improve the robot's SA based on the human behavior models. Implementing knowledge about these specific behaviors of the users with respect to several modalities will help the robot to interpret and better understand the actions. In other words, knowledge about task-related user behavior can help to improve the robot's SA.

In future work, more modalities that the users employ to interact with the robot shall be integrated in the human behavior models (for example, speech and orientation of the head). Moreover, the gained knowledge cannot only be used to improve the robot's perception of the interaction but it can also be exploited to adapt the robot's own behavior to the human behavior models. In this sense, the results can help to improve HRI in general.

REFERENCES

[1] L. A. Suchman. *Human-Machine Reconfigurations: Plans and situated actions*. 2nd edition. Cambridge: Cambridge University Press (2007).

[2] M. R. Endsley. Design and Evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society 32nd Annual Meeting.* Santa Monica, CA: Human Factors Society, 97-101 (1988).

[3] M. Lohse. *Investigating the influence of situations and expectations on user behavior - empirical analyses in human-robot interaction*. Doctoral Thesis. Bielefeld University, Technical Faculty, Germany (submitted).

[4] J. T. Hackos and J. C. Redish. *User and task analysis for interface design*. New York: Wiley and Sons (1998).

[5] A. Shepherd. *Hierarchical Task Analysis.* London and New York: Taylor and Francis (2001).

[6] J. A. Adams. Human-Robot Interaction Design: Understanding User Needs and Requirements. *Proceedings of the 2005 Human Factors and Ergonomics Society 49th Annual Meeting (2005).*

[7] K. Severinson-Eklundh, A. Green, H. Hüttenrauch, L. Oestreicher, and M. Norman. Involving Users in the Design of a Mobile Office Robot. *IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews 34(2),* 113-124 (2003).

[8] H. Kim and D. Kwon. Task Modeling for Intelligent Service Robot using Hierarchical Task Analysis. *Proceedings of the 2004 FIRA Robot World Congress,* Busan, Korea (2004).

[9] M. Argyle. *Bodily Communication*. London: Methuen (1988).

[10] J. Aiello. Human spatial behavior. In D. Stokols and I. Altman (Eds.). *Handbook of Environmental Psychology 1*. New York: Wiley Interscience, 505-535 (1987).

[11] L. Hayduk. Personal space: Where we now stand. *Psychological Bulletin, 94*, 293-335 (1983).

[12] Website of the Cogniron project. www.cogniron.org, 24. September 2009.

[13] D. S. Syrdal, M. L. Walters, K. L. Koay, and K. Dautenhahn. The role of autonomy and interaction type on spatial comfort in an HRI scenario. *Proceedings of "Robotic Helpers: User Interaction, Interfaces and Companions in Assistive and Therapy Robotics", Workshop at ACM/IEEE HRI 2008,* Amsterdam, the Netherlands, 12 March 2008, Technical Report 470 University of Hertfordshire, UK (2008).

[14] M. L. Walters, K. L Koay, S. Woods, D. S. Syrdal, and K. Dautenhahn. Robot to Human Approaches: Preliminary Results on Comfortable Distances and Preferences. *Technical Report of the AAAI Spring Symposium on Multidisciplinary Collaboration for Socially Assistive Robotics, (AAAI SS07-2007),* Stanford University, Palo Alto, CA, USA (2007).

[15] S. Suzuki, Y. Morishima, M. Nakamura, N. Tsukidate, and H. Takeda. Influence of body orientation and location of an embodied agent to a user. *Proceedings of the 20th International Conference on Computer Animation and Social Agents (CASA2007),* Hasselt University, Belgium, 1-10 (2007).

[16] A. Kendon. *Conducting Interaction – Patterns of Behavior in Focused Encounters.* Cambridge, MA: Cambridge University Press (1990).

[17] H. Hüttenrauch, K. Severinson-Eklundh, A. Green, and E. A. Topp. Investigating Spatial Relationships in Human-Robot Interaction. *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems,* Beijing, China (2006).

[18] M. Argyle. *Social Interaction*. London: Methuen & Co Ltd (1969).

[19] Y. Nagai, M. Asada, and K. Hosoda. Learning for joint attention helped by functional development. *Advanced Robotics, 20 (10)*, 1165 – 1181 (2006).

[20] M. Staudte and M. Crocker. The utility of gaze in spoken human-robot interaction. *Proceedings of Workshop on Metrics for Human-Robot Interaction 2008, March 12th,* Amsterdam, the Netherlands, 53-59 (2008).

[21] C. Breazeal, C. D. Kidd, A. L. Thomaz, A. L.; G. Hoffman,and M. Berlin, Effects of Nonverbal Communication on Efficiency and Robustness in Human-Robot Teamwork. *Proceedings of International Conference on Intelligent Robots and Systems (IROS)* (2005).

[22] C. L. Sidner, C. Kidd, C. Lee, and N. Lesh. Where to look: A study of human-robot engagement. *In Intelligent User Interfaces (IUI),* Funchal, Island of Madeira, Portugal, January 2004, 78-84 (2004).

[23] A. Green. *Designing and Evaluating Human-Robot Communication. Informing Design through Analysis of User Interaction.* Doctoral Thesis. KTH Computer Science and Communication, Stockholm, Sweden (2009).

[24] J. E. Young, R. Hawkins, E. Sharlin, and T. Igarashi. Toward Acceptable Domestic Robots: Applying Insights from Social Psychology. *Int. Journal of Social Robotics, 1,* 95-108 (2009).

[25] M. Hanheide and G. Sagerer. Active Memory-based Interaction Strategies for Learning-enabling Behaviors. *Proc. International Symposium on Robot and Human Interactive Communication (RO-MAN'08)* (2008).

[26] E. Hall. A System for the Notation of Proxemic Behavior. *American Anthropologist, 65*, 1003-1026 (1963).

[27] Elan annotation tool. http://www.lat-mpi.eu/tools/elan