The role of expectations in HRI

Manja Lohse[1]

Abstract. Whenever humans interact they form expectations which their behaviour is grounded in. This paper claims that this is also true for human-robot interaction (HRI). It introduces assumptions from expectation theory and illustrates how these can be tested in an HRI context. The paper aims to show that knowing the users' expectations can help to design robots. Since user expectations depend on the robot behaviour, they can be shaped to improve the interaction.

1 INTRODUCTION

The term "expectation" is used in many contexts in every-day life: employees have expectations toward their employers, the society has expectations toward politicians, and someone can expect a baby. Formally, expectations refer to norms, as well as to what happens in the case of uncertainty. Norms are general design guidelines which strongly depend on cultural factors, e.g., distances a robot has to keep from the human (e.g. [1],[2]). Within a culture these norms are rather stable. In contrast, expectations in case of uncertainty strongly depend on the person and on the situation. Based on their own experiences, all people form their own expectations. These are driven by conscious and unconscious memories.

Concerning norms (prior expectations based on culture that users bring into the interaction situation) some studies have been conducted in the context of HRI. These do not use the term expectations but rather refer to attitudes. One study has been reported by Khan already in 1998 [3] when even fewer people than today had interacted with a robot. The findings of the survey with 134 Swedish subjects suggested that certain tasks were preferred for intelligent service robots: robots should help with chores like polishing windows, cleaning ceilings and walls, cleaning, laundry, and moving heavy things. However, they were not wanted for tasks like baby sitting, watching a pet, or reading aloud. 69% of the subjects found the idea of a robot as assistant positive, 44% found it probable, 23% frightening, and 76% found the concept useful [3].

In 2006, Arras and Cerqui [4] have supported the idea that the attitude towards robots depends on the tasks they are built to accomplish. In a survey with more than 2000 participants at the Swiss expo, 71% of the interviewees exhibited a neutral attitude towards robots; meaning that they neither rejected them nor were completely positive. This was due to the fact that the result was based on the mean of attitudes toward robots used for different tasks. As mentioned above, some tasks were very welcome while others were not.

These findings have recently been supported by Ray and colleagues [5] in their publication titled "What do people expect from robots?". In a questionnaire study on people's perception of robots with 240 participants, the authors found a positive attitude towards robots in general. But they also reported that people prefer robots mainly as helpers in daily life and not as caregivers in whatever context. They noted that people expressed their fear about robots replacing humans and about robots becoming too autonomous. The authors' findings revealed that robots should not be humanoids or look like living beings. On the other hand, while appearance should not be human-like, people prefer natural channels, e.g., speech, for communication with robots.

Attitudes have also been in the focus of research with the NARS (negative attitudes towards robots) scale, which is based on the concept of computer anxiety (e.g. [6], [7]). However, researchers using this scale barely identified a link between attitudes and behaviour. In this sense, knowing one's attitude does not allow for a reliable prediction of actions because the behaviour is also strongly influenced by the situation [8]. In general, a positive attitude towards robots cannot lead straight to the conclusion that people in fact use the systems. Therefore, I state that a priori expectations and attitudes are not the main determinants of individual interactions. Instead, it is the situation that mainly affects the expectations during the interaction. Just as situations, in my approach expectations are seen as dynamic. This assumption is based on a constructivist model of mind (see Section 2B). Similar to memories, expectations need not only be recalled but they are actively constructed in a situation. Thus, the situation strongly influences what comes to a person's mind (accessibility of an expectation), which schemas are used for interpretation explicitly or implicitly (explicitness of expectations), how a person interprets what is occurring and which attribution processes take place (how the person logically justifies the occurrences). Therefore, I assume that changes of expectations can be observed during the interaction. This approach attempts a more reliable explanation of user behaviour in HRI than attitudes and a priori, normative expectations because it generalizes less and takes into account the particularities of the situation and online changes of expectations during the interaction.

The first part of the paper introduces expectation theory, revealing different kinds of expectations (Section 2). Section 3 provides some thoughts on the importance of the situation in which an interaction is taking place. In Section 4 assumptions based on the theory are summed up. Thereafter, I present a case study which shows how the assumptions might be tested based on data from user studies (Section 6). The data used for the analysis is described in Section 5. Based on the analysis, I discuss the relevance of expectations in HRI (Section 7).

---

[1] Applied Informatics, Research Institute for Cognition and Robotics, Bielefeld University, Bielefeld Germany.

Email: `mlohse@techfak.uni-bielefeld.de.`

2 EXPETATION THEORY

Before taking a closer look at expectation theory, I want to clarify why in this paper both the terms "expectation" as well as "expectancy" are used. I agree with [9] that in the literature it seems that both termini appear interchangeably. I prefer "expectation" but if literature is cited, I use the same term as the authors. The discussion of a possible distinction between the terms is not in the scope of the current paper.

In a recent paper, Roese and Sherman [10], give an overview of expectancy theory. They define expectancies as follows:

> "Expectancies are beliefs about a future state of affairs, subjective estimates of the likelihood of future events ranging from merely possible to virtually certain." (p.91)
>
> "The expectancy is where past and future meet to drive present behavior" (p.92)

In the first version of the article from 1996, Olson, Roese, and Zanna [11] call the kind of expectationcies defined here *probabilistic expectancies* since they describe what could happen with a certain probability. The definition presented above indicates that past experiences shape expectancies. On the other hand, expectancies support anticipating the future and, therefore, shape the behaviour of a person.

One classic experiment in social psychology which shows the importance of expectations was conducted by Kelley (1950) (see [12], [13]). In the experiment, students were told that a guest instructor was either warm or cold. After the lecture students judged the instructor. Their judgments clearly showed assimilation effects: the group of students that were told that the instructor was warm hearted, rated him as being more considerate, informal, sociable, popular, good-natured, humorous, and humane. These students were also more open for a discussion with the instructor. In this example expectations primed in the students beforehand influenced information processing. Based on this experiment one could assume that participants enter situation with expectations which do not change over time. In contrast to the experiment in most situations, and also in our experiments, however, people are not primed for specific expectations beforehand. How expectations form if they are not primed is described in the following section.

A Sources of Expectations

Olson and colleagues [11] name three major sources of expectations: direct personal experience, communication from other people (indirect experience), and beliefs that are inferred from other beliefs. Personal experience determines whether expectations are *semantic* or *episodic*. According to Roese and Sherman [10], semantic expectancies are

> "preexisting knowledge structures that are extracted from ongoing experience, stored in memory, and retrieved when needed. By contrast episodic memories must be formulated on the spot before they can be applied" (p.95).

While semantic expectancies are an efficient way of using past experiences, episodic memories provide depth by not including abstract knowledge but information from the situation. Semantic and episodic expectancies interact and can be applied in parallel. Usually, people at the outset have specific expectations, which become more general over time, and are finally balanced at mid-levels. Based on this concept, one can assume that expectations change over time. Especially in situations in which users interact with a robot for the first time changes are probable.

In 1977, Heckhausen [14] has introduced a model that also points out the influence of the situation on expectancies. The model encompasses a four-stage sequence of events:

- the initial situation as perceived
- the person's action
- the outcome of the action or the situation
- the consequences of the outcome

Based on these stages, Heckhausen names different kinds of expectancie: situation-outcome expectancies, action-outcome expectancies (these two are typically addressed in motivation theory, e.g., by Tolman), action-by-situation-outcome expectancies, and outcome- consequence expectancies.

Maddux [15] applies two different terms (behaviour outcome expectancy and stimulus outcome expectancy) which seem to be similar to action-outcome expectancies and situation-outcome expectancies, respectively. Behaviour outcome expectancies are beliefs that a specific behaviour results in a specific outcome or set of outcomes. Stimulus outcome expectancies "are concerned with beliefs that certain events provide a signal or cue for the possible occurrence of other events" ([15], p.23).

The next section provides some more details about the function of expectations which is closely connected to their processing.

B Function and Processing of Expectations

> " […] people think about the future and use such thoughts in ongoing judgment, reasoning, decision making, and behavior." ([10], p.91)

This statement shows that expectancies influence behaviour. They provide an effective way to guide behaviour by providing a shortcut in mental processing. Judgments about the future are not stable but rather construed by the human. In the construction process expectations guide information gathering as top-down processes influencing event processing; and they provide structure and meaning for the interpretation of the gained information. On the other hand, they can also lead to inaccurate inferences if they are incorrect, if they bias information collection, or if they overrule consideration of information altogether [16].

In accordance with these findings, Roese and Sherman [10] state that expectations influence information seeking. Sometimes people try to validate their hypothesis of the world (at least for subjective expectancies) and search for more information. Validation processes occur especially in cases of expectation disconfirmation. Disconfirmation is identified with a mechanism that Roese and Sherman term regulatory feedback loops. According to the authors, the current state of a situation is compared to an ideal or expected state. If discrepancies are discovered, the person will search for behaviour in order to correct or reduce the discrepancies.

> "Behavior control therefore requires continuous online processing in the form of continuous comparison, or pattern matching, between the current state and the expected state. Very likely in parallel to this comparative process is the conceptually similar online comparison between the current state and recent past states." ([10], p.93)

In most cases expectations are confirmed since behaviour is constantly based on simple expectations, e.g., if we push a light-switch we expect that the light comes on (or goes off if it was on beforehand) because we have experienced this many times before. The confirmed expectations are usually processed heuristically. Then again, heuristically processed information is more likely not to bring to awareness disconfirmation. Disconfirmation of expectations causes more careful processing. The expectation is perceived as being less certain. Disconfirmed expectations become more explicit and accessible and are therefore more likely to be tested in the future. Next to these basic assumptions of expectation disconfirmation, Olson et al. [11] list some more cognitive consequences of disconfirmation:

- both consistent and inconsistent information receive more processing than irrelevant information
- people prefer to interpret information in line with their expectancies
- unexpected events trigger attributional processing but attributions also influence expectancies (reciprocal relationship)
- disconfirmation triggers more counterfactual processing, but expectancies also shape the semantic content of counterfactual thought (reciprocal relationship)
- expectancies provide a cognitive structure to more easily encode and retrieve consistent information from memory, on the other hand by means of expectancies unexpected events become salient which increases memorability

Olson et al. [11] conclude that behaviour is consistent with the content of expectancies, e.g., people chose tasks in which they expect to succeed and put more effort into them. Expectations can also lead to self-fulfilling prophecies (e.g., [11]). Next to behavioural also affective and physiological consequences (e.g. placebo effect) can be found. Affect usually is negative when expectancies are disconfirmed an exception being, e.g., a surprise party where the expectations are disconfirmed in a positive way.

3 IMPACT OF THE SITUATION

In Section 2A, the model by Heckhausen which links expectations and situation has been introduced. Moreover, it has been discussed how expectations are construed in the situation and how the situation determines what comes to the mind of a person. In this section, I aim at further clarifying how the term situation is understood in this paper.

In contrast to personality traits, situation is a dynamic concept [8]. Many experiments have shown how situational factors superpose personality in human-human interaction; e.g., in Libermann, Samuels and Ross (2004) (in [17]) students played the same game which was introduced to two groups with two different names (Community Game and Wall Street Game). Students in the second condition (Wall Street Game) displayed significantly more competitiveness independent of how their competitiveness was rated beforehand. They obviously constructed a certain idea of the situation. The awareness of the situation was affected by the name of the game.

Situation awareness (SA) as a concept has been formulated by Endsley and colleagues [18]:

> "The formal definition of SA is 'the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future'" (p.13; from Endsley, 1988)

Endsley differentiates three levels of situation awareness: perception (Level 1), comprehension (Level 2), and projection of future states (Level 3). The definition shows that SA is closely connected to expectations in that the person analyzes the surroundings to predict future occurrences and to adjust the behaviour. According to Endsley, expectations, next to goals, mental models, and schemas form the internal representation of state. They influence perception, attention, and other cognitive processes. These, in turn, influence expectations. Therefore, expectations are bound to the situation and to how it is perceived.

In the example by Liberman and colleagues introduced above, the researchers supposed that all subjects conceptualized the situation in the same way based on the name of the game that they played. This is not necessarily true for HRI situations. Fischer [19] has analyzed experimental data for cues how the subjects conceptualized situations. She found that beliefs about the communication partner show in linguistic behaviour and situation influences the language used.

> „the speakers have been found to constantly define and, if necessary, redefine what they understand the situation to consist in, depending on their current hypotheses about their communication partners and on their own emotional state." ([19], p.7)

This work emphasizes, once again, the importance of the situation as such. However, while Fischer is interested in linguistic phenomena, I focus on strategies that include different modalities and what they tell us not only about the perception of the situation but also about users' expectations. Since processing is an online mechanism, I expect that expectations change during interaction, and that the changes are based on what the robot does and what happens in the situation as such.

4 ASSUMPTIONS BASED ON EXPECTATION THEORY

Summarizing the previous sections, I assume that:

1. expectations are construed in the situation and, therefore, may change during the interaction
2. expectations are influenced by the robot behaviour, and the situation.

3. expectations are influenced by the human characteristics (e.g., appearance, age)
4. first contact situation results in not having concrete former expectations → people form episodic expectations and repeat successful strategies
5. information gathering is based on expectations, what people pay attention to (look for, ask) displays what they expect
6. users try to verify their expectations especially after expectation disconfirmation
7. expectation disconfirmation can lead to changes in affect
8. behaviour is consistent with the content of expectations, e.g., people chose tasks in which they expect to succeed and put more effort into them

These assumptions provide an extensive framework for research on expectations in HRI. Not all of them can be tested in this paper. However, all the assumptions will have to be tested if we want to show the influence of expectations on HRI. In the following I introduce a case study as an example on how to determine expectations based on a data-driven approach.

5 HRI DATA

The relation between HRI and expectations can only be researched based on interaction data which is situated in a certain situation because (1) real users can be expected to have different expectations and (2) as has been discussed before, the situation determines the expectations. The test bed in this paper is data from a study in a home tour situation. The home tour focuses on multi-modal human-robot interaction in which the robot should learn about the domestic environment and its artefacts, the appearance and location of objects, and their spatial-temporal relations. With these abilities the robot can adapt to new settings like a user's home. The home tour focuses on the tutoring of the robot. The exploitation of the acquired knowledge and representations for, e.g., conducting fetch-and-carry tasks is a consequent extension of this scenario. Robots' capabilities for natural interaction in the home tour comprise understanding of spoken utterances, co- verbal deictic reference, verbal output, referential feedback, and attention to persons. Furthermore, the robot must be able to follow a human when she is walking around and showing an apartment.

Robot platform

The robot used for the study is called BIRON (BIelefeld Robot companiON). BIRON is based on a Pioneer PeopleBot platform. As can be seen in Fig. 1, BIRON employs various modalities: it can see faces with the camera on top and gestures with the camera below the screen, it displays a character called Mindi on the screen (see Fig. 2, 3, and 4), has microphones for speaker localization (speech recognition is done using a headset), speakers for speech output, and a laser scanner to explore the environment. Details about architecture, design principles and abilities of BIRON can be found in [20].

Subjects

14 subjects participated in the user study. All participants were German native speakers and interacted with BIRON in German. They were recruited at a public event of the university. Their age ranged from 16 to 71 years (average 45.5 years). 9 participants were male, 5 female.

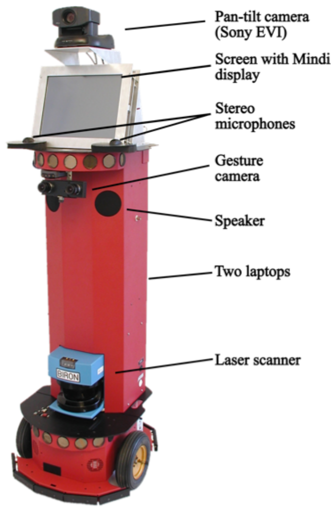

Fig. 1 The robot BIRON

Procedure

On arriving, the participants were welcomed and introduced to the study. They answered a questionnaire on demographic data and their experience interacting with robots. Afterwards, the subjects were trained on using the speech recognition system, i.e., they were instructed about the proper placement of the head set microphone and were asked to speak some phrases for habituation. The recognition results were displayed in verbatim on a laptop computer. Thereafter, participants were guided into the room where the robot was waiting ready for operation. Subjects were assisted during the first contact in order to reduce hesitant behaviours. They were handed a tutorial script for practise. The script contained all commands the users would need later on and the autonomous robot operated just as it did after the training session. During this initial tutorial session the experiment leader also instructed the users on how to pull the robot to alleviate cases when the robot gets stuck. After the tutorial session the training script was handed back and the main task, namely to teach areas and locations in several rooms, was carried out by the subjects. The instruction for this main task was:

- guide the robot through the apartment, i.e. from the living room to the dining room via the hall
- show and label the living room and the dining room
- show the cupboard in the living room and the floor lamp in the dining room

During this part of the interaction, the experiment leader only intervened when prompted by the subjects. The whole interaction was videotaped.

Afterwards, the subjects were interviewed about the interaction in general and the visual display in particular. The interviews have also been digitized. Moreover, the participants answered a second questionnaire which included items on liking of the robot, attributions made towards the robot and usability of the robot. More information about the study is provided in [21].

Data Analysis

The video data of the study was used for the analysis presented here. The corpus was annotated using ELAN. A coding scheme for human speech, gestures, and other movements was developed and employed by three coders. The robot activities were logged automatically. All data (video, manual annotations, and automatic logs) were synchronized in one ELAN file per user in (see [21] for more information).

For the analysis presented here, one situation was chosen from all recordings. The situation is characterized by BIRON having an incomplete percept of the user. The percept consists of a voice and visual features, namely legs and face. If the person perception is unreliable, e.g., because lightening conditions are poor or because the person has not said something in a while, the robot is not able to match the percept to one person.

This state is communicated on the display with a picture of a the screen character Mindi (Fig. 2). Due to the state the robot does not react to the user until the she has done something to improve the perception. In general, solutions to this problem include two strategies: verbal behaviour or movement of the user in front of the robot. The user has the initiative. Since this state is not part of the explicit dialog, it is only communicated using the Mindi display. The situation perception of the users and their expectations determine how they react to the situation. Therefore, the situation is of special interest here.

The situation was chosen because it was an unexpected event in the interaction. In fact, it constitutes a change in the interaction that should also lead the user to adapt her expectations to the new situation. The modified expectation should then result in a change in user behaviour.

In the video data of the 14 subjects, 26 sequences were identified. In the interactions of users 10 and 13 no such sequences occurred. All other recordings contained one to three sequences.

In the most basic form, the sequences are characterized by the following actions. First, the Mindi display changes to "poor person perception" (Fig. 2), the person reacts to the display, the display changes to the "processing" Mindi (Fig. 3), then to the "robot is facing user" Mindi (Fig. 4), and finally the robot says "hello". The average length of these sequences is 13.8 seconds (minimum 4 seconds, maximum 45 seconds; measured from Mindi appearing to Mindi changing to processing).

Apart from the video recordings, the questionnaire that the subjects answered after the interaction was taken into account for the analysis. The questions of interest are related to whether the participants had seen the "poor person perception" Mindi (Fig. 2) and what they thought it meant.

Another aspect that was included in the analysis was that five subjects had received an introduction to the meaning of the picture during the interaction. In the explanation BIRON told them: "If you see this picture, I cannot see or hear you very well. Please change your position and repeat what you have said." Only five people listened to the explanation because the subjects were free to chose whether they wanted to listen to it or not. In Section 7, it will be discussed whether receiving an introduction or not had an impact on the interaction.

Moreover, users 5 and 6 received help from the experimenter when the Mindi in question was displayed on the screen during the training phase. User 5 was told to restart the robot (by saying "restart"). User 6 was asked to say "hello" again in one occasion and to restart the robot in another.

In the training the experimenter at times interrupted the interaction and the subjects often continued with the training script regardless the reaction of the robot. Therefore, only sequences from the main task were included in the analysis.

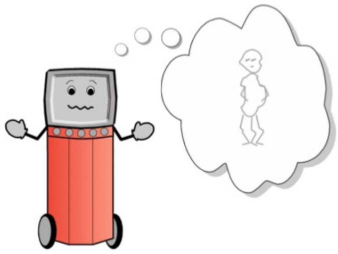

Fig. 2 "Poor person perception" Mindi

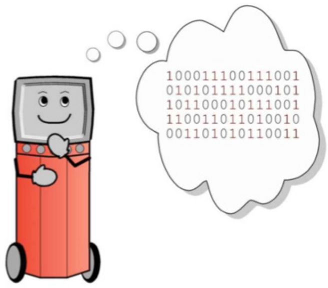

Fig. 3 "Processing" Mindi

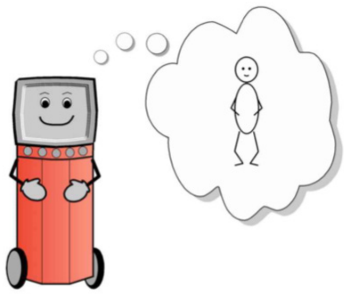

Fig. 4 "Robot is facing user" Mindi

6 RESULTS

I want to anticipate that in fact the new situation changed the expectations of the users. However, there were differences in how they conceptualized the change.

One aspect that all analyzed sequences had in common was that the users searched a position in front of the robot before starting another behaviour to resolve the situation. Users did not only aim to stand in front of the robot body but also tried to be faced by the camera on top of the robot. This could be seen in the videos because all users kept walking around and leaning to the side until the robot faced them. Therefore, the poor perception display triggered the same expectation in all user that the robot could perceive them better when facing them.

Another behaviour that was shared by all participants was that the time users wait before the next action (verbal or movement) was very stable within subjects. However, it strongly differed between subjects. Some only waited two seconds while others waited ten seconds and more for BIRON to give a feedback. The expected feedback could be verbal but also appear on the screen.

The results clearly showed that when the Mindi display changed to "processing", the users did not take any more actions and waited for more robot feedback.

Another expectation that was shared by all participants was that some verbal utterance was necessary to regain the robot's attention. However, three different strategies to resolve the situation could be identified:

1) verbal behaviour only
2) verbal behaviour first with movement added
3) movement first with verbal behaviour added

Their distribution is depicted in the following table.

Table 1 Strategies applied by subjects

user	1	2	3	4	5	6	7	8	9	11	12	14
strategy 1	3	3	1	2		1		2	1		2	
strategy 2						1	1	1			1	1
strategy 3				1	3				1	1		

Strategy 1 was characterized by verbal behaviour only. Out of the 26 analyzed sequences, 13 were identified with this group. They mainly occurred in short sequences in which it sufficed that users said something once (either "hello" or something that starts another action like "follow me"); e.g. as depicted in Fig. 5 (time code in the figure caption tells when the sequence occurred in the interaction). The figure illustrates occurrences over time. In the upper row actions of the human are labelled, in the lower row actions of the robot. Red (dark colour) actions are verbal actions while light-grey actions are non-verbal actions (movements of the human, changes on BIRON display).

The average length of these sequences was 7.2 seconds (minimum 4 seconds, maximum 16.5 seconds). In general the sequences were shorter than 8 seconds. The average time people waited after the last action was 6.2 seconds (minimum 3.25 seconds, maximum 7.5 seconds). Measured from the appearance of the Mindi it was 4 seconds (minimum 1.5 seconds, maximum 6 seconds). All these numbers do not include user 12 because she took significantly longer than all other subjects and was obviously distracted by some event that was not connected to the interaction.

In addition to the sequences which were resolved after giving a verbal command once, this group also includes two situations (two different subjects), in which BIRON did not react after the first utterance and the subject only repeated the verbal utterance without any movement.

Strategy 2 contains five cases of five users all of which apart from this strategy used strategy 1 in other sequences. The rest of their sequences were short sequences as described above (saying "hello" or another command once and the interaction continued). In strategy 2, sequences contained saying "hello BIRON" or "BIRON hello", movement, and at least one more verbal command (Fig. 6). Interestingly, the movements of the users almost only consisted of stepping backwards and forwards in front of the robot (4 instances). Only one person stepped to the left or the right (1 instance). Moreover, in all sequences the exact same wording was repeated. The average length of sequences in this group was 18.25 seconds (minimum 8.75 seconds, maximum 24 seconds).

Strategy 3 includes six cases, three of which by user 5 (all her sequences) and one each of three other users. The strategy was characterized by the user moving after positioning herself in front of the robot and before taking any verbal action.

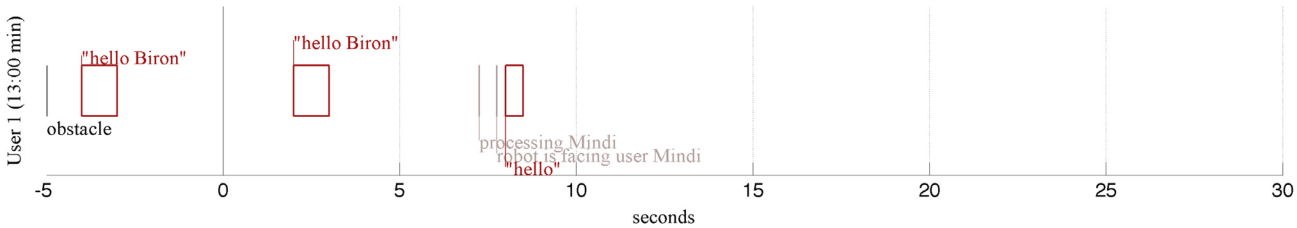

Fig. 5 User 1 (13:00 min)

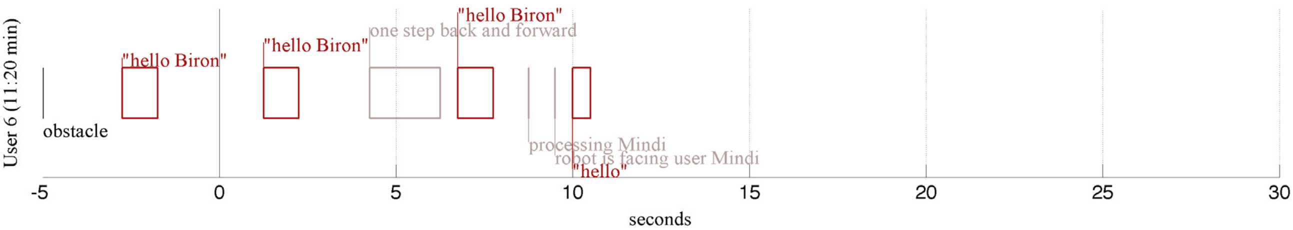

Fig. 6 User 6 (11:20 min)

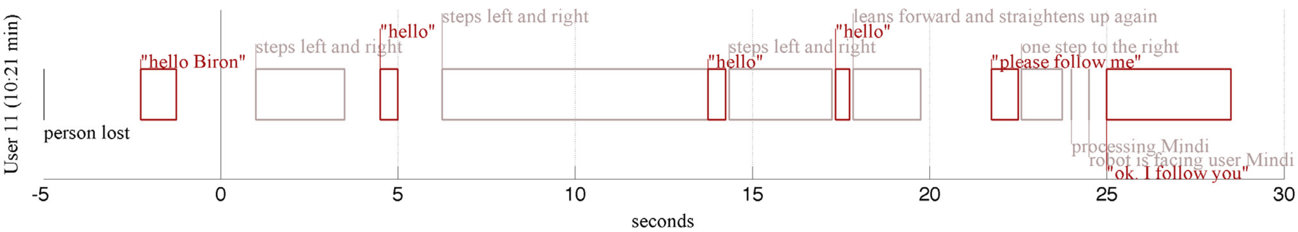

Fig. 7 User 11 (10:21 min)

Altogether, the pauses between movements were shorter than between speech. User 11 actually kept moving almost continuously with only four breaks in a sequence of 24 seconds (Fig. 7).

The average length of sequences with this strategy was 26.5 seconds (minimum 8.25 seconds, maximum 45 seconds) which is much longer than in Strategy 1 and 2.

In contrast to strategy 2, movement to the side is more common here (to the side - 20 times, backward/forward - 2 times, legs apart - 2, times lean forward - 3 times). Only one person in one sequence stepped forward and backward (user 5, 14:06). User 4 (twice in one sequence) and 11 (once) lean forward and position their faces in front of the camera.

Next to the video data we also analyzed the questionnaires to find out why different strategies were employed. Of interest here is how the users interpreted the Mindi in question (Fig. 2). They were asked this as an open question ("What does this picture express?) which they answered in writing. The following descriptions were given (answers were translated by the author):

User 1	BIRON has not understood "hello", is waiting for commands
User 2	BIRON cannot see / hardly hear the user
User 3	BIRON does not see
User 4	BIRON can hardly see the user
User 5	BIRON tries to recognize the user
User 6	BIRON has lost sight of the user
User 7	BIRON does not know / has not understood
User 8	BIRON does not find the person
User 9	the user is unknown/ not recognized any more
User 11	BIRON does not know
User 12	BIRON does not see / does not know what to do with a command
User 14	BIRON does not know what to do in a situation

While nine answers (1, 2, 3, 4, 5, 6, 8, 9, (12)) actually reflected poor perception of the users by the robot, others referred to knowledge and understanding of the robot (7, 11, (12), 14). In most cases, user perception addressed the robot's ability to see (apart from users 1, 2, 7). The answers of users 10 and 13 are not included because their video data did not contain sequences that were analyzed.

The question which arises from the different strategies that can be seen in the videos and the answers from the questionnaires is how users conceptualized the situation and what expectations they formed. It will be discussed in the following section.

7 DISCUSSION

In the following, I want to elaborate on how the findings from video data and questionnaires can help to test assumptions regarding expectations in HRI.

The results showed that, in the situation analyzed here, it was crucial that all users felt that the robot's camera was facing them. This expectation arose in the situation based on the robot's screen output and missing verbal feedback. Based on their experience users new that something went wrong in the interaction. All users reacted to the display with verbal and non-verbal behaviour which signalled that they took over initiative and asked for the robot's attention. Thus, the robot feedback triggered the expectation that the user had to take the turn.

In the sequences analyzed, it was noticeable that all participants accepted the Mindi changing to processing as feedback signal. As soon as the display switched, the users then expected the robot to have the initiative and did not take any more action. Therefore, the display seems to indicate that the robot is doing something and some action will follow. On the other hand, for the poor person perception Mindi the expectation did not seem to be quite clear and users conceptualize the situation differently. This became obvious in two observations: the time that users waited before the next action was quite stable within subjects while the time between users differed; and participants' behaviour in this situation can be grouped in three strategies (verbal behaviour only, verbal behaviour enriched with movement, and movement enriched with speech).

Strategy 1 proved to be the most successful one in that the sequences were shortest and the problem was resolved fastest. Verbal input by the user allowed the robot to identify the interaction partner and to continue with the task. In accordance with expectation theory, users repeated this successful strategy when the same situation reoccurred. Only two people gave the robot verbal feedback also if it failed at first. Their expectation seemed to be that BIRON needed a verbal percept which was probably furthered by many requests of the robot to say "hello" if something went wrong (e.g. after moving the robot away from an obstacle, after having lost the percept of the person completely).

Users that applied strategy 2 also seemed to agree with the importance of greeting the robot to attract its attention even though the questionnaire data showed that their understanding of the Mindi picture differed. Only user 2 did not greet BIRON but started a new action (following). He also was the only one who moved during the utterance, all others only tried one action (speak, move) at a time. His movement was closely connected to the follow command. Therefore, with reference to the situational constraints, the function seems to be to start walking in the right direction rather than to attract attention. What is common to strategy 2 is that verbal utterances were repeated and enriched with movement. The direction of the movement in most cases was forward and backward.

In contrast people applying strategy 3 moved to the left and to the right. This movement was accompanied by a camera movement to both sides. Therefore, it was an obvious feedback that the robot was still doing something. This might have triggered subjects to try out more movements. No camera movement occurred when people walked backward and forward. Therefore, in sum, people applying strategy 2 might have tried more verbal behaviour.

Non-verbal behaviours in strategy 3 also included that user 4 (twice in one sequence) and 11 (once) leaned forward and positioned their faces in front of the camera. They probably assumed that the robot had difficulties to perceive the face but needed this percept in order to continue interaction. This movement was not tried out right away but rather late in the sequence of user 11 and only in the third sequence of user 4.

Strategy 3 was usually only used once. Only user 5 repeated the strategy which was not successful with regards to the time needed to resolve the situation. However, it is noticeable that in the later situations she used a verbal utterance much earlier after the Mindi appeared. Obviously her expectations had changed at least in part. This finding supports the assumption, that the history of the interaction influences expectations.

Concerning the question why people apply a certain strategy, one possible explanation is the training participants received from the robot itself and hints from the experimenter which influence their perception of the situation. As described in Section 5, five subjects had received the introduction to the meaning of the Mindi picture during the interaction. Their behaviour was different to the other subjects in that five of six cases of strategy 3 were used by people who received BIRON's instructions. The verbal introduction by BIRON seems to stress more strongly that the robot has a problem "seeing" than the Mindi picture does. However, it is noticeable that user 11, who received the introduction, interpreted the picture as "BIRON does not know", even though there was another picture with a question mark in the speech bubble which was clearly recognized as BIRON does not know / does not understand.

These finding support the assumption that expectations are influenced by the robot's behaviour. Therefore, robot behaviour should be designed to shape users' expectations and behaviour to enable them to more efficiently solve tasks in the interaction. In the data reported here such an improvement regarding efficiency could not be observed because (1) as the questionnaire entry of user 11 shows users cannot be expected to perceive all information equally well, (2) too few situations were analyzed, and (3) even if more data would have been available it can be doubted whether it would have revealed the improvement because, as the analysis has shown, the feedback in this situation is obviously not yet clear enough to shape users' expectations. However, now having analyzed the data, we know about the implications the robot behaviour has on user experience and can strive for a more helpful feedback. The example, moreover, shows that data acquired in user studies can be used to infer expectations that users have during the interaction. Moreover, it can be assumed that these expectations change based on the situation and on how it is conceptualized by the user.

8 Future Work

This paper is only a start in research regarding expectations in HRI. It introduces assumptions on user expectations which need to be tested to fully explore the influence of expectations on HRI. Based on a case study, it claims that expectations can be inferred from data acquired in interaction situations and that interpretation of data can profit from taking users' expectations and conceptualizations of the situation into account. Knowing the users' expectations can help to design robots because expectations depend on the robot behaviour and can therefore be shaped to improve the interaction.

Another approach that in the future might drive robot design is to not only research user expectations but to also explore robot expectations. Answering the questions "What kind of input does the robot expect?" and "How can it communicate which input is expected?" might also improve the interaction considerably.

Finally, the question on how expectations can be measured more reliably deserves further consideration. The case study presented here revealed some expectations in a very restricted situation. Since these findings should be compared to other interaction situations all of them should be using the same measurements.

REFERENCES

[1] E. Pacchierotti, H.I. Christensen, and P. Jensfelt. Human-Robot Embodied Interaction in Hallway Settings: a Pilot User Study. ROMAN 2005. pp. 164-171 (2005).

[2] M.L. Walters, K. Dautenhahn, K.L. Koay, C. Kaouri, R. Te Boekhorst, C. Nehaniv, I. Werry, D. Lee. Close encounters: Spatial distances between people and a robot of mechanistic appearance. In: Proceedings of the IEEE-RAS International Conference on Humanoid Robots (2005).

[3] Z. Khan, Attitudes towards intelligent service robots. Technical Report Nr. TRITA-NA_P9821, IP-Lab-154 IPLab, Royal Institute of Technology (KTH), Sweden (1998).

[4] K.O. Arras and D. Cerqui. Do we want to share our lives and bodies with robots? A 2000-people survey. Technical Report Nr. 0605-001 Autonomous Systems LabSwiss Federal Institute of Technology, EPFL (2005).

[5] C. Ray, F. Mondada, and R. Siegwart. What do people expect from robots? *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nice, France. pp. 3816-3821 (2008).

[6] T. Nomura, T. Kanda, and T. Suzuki. Experimental Investigation into Influence of Negative Attitudes toward Robots on Human-Robot Interaction. Presented at the *3rd Workshop on Social Intelligence Design (SID2004)*, Twente (2004).

[7] T. Nomura, T. Kanda, T. Suzuki, and K. Kato. Psychology in Human Robot Communication: An Attempt through Investigation of Negative Attitudes and Anxiety toward Robots. *RO-MAN 04* (2004).

[8] K.J. Rohlfing, M. Rehm, and K.U. Goecke. Situatedness: The interplay between context(s) and situation. *Journal of Cognition and Culture* 3(2) pp. 132-157 (2003).

[9] N. T. Feather. Introduction and Overview in: N. T. Feather (ed.) *Expectations and Actions: expectancy-value models in psychology.* Lawrence Earlbaum Associates: Hillsdale. pp. 1-14 (1982).

[10] N.J. Roese and J. W. Sherman. Expectancy. In: A. W. Kruglanski and E.T. Higgins, (ed.) *Social Psychology. Handbook of Basic Principles.* Guilford Press. pp. 91-115 (2007).

[11] J. Olson and N. Roese and M. Zanna. Expectancies. in E.T. Higgins and A. Kruglanski (ed.) *Social Psychology: Handbook of Basic Principles*. Guilford Press: New York. pp. 211-238 (1996).

[12] M. Biernat. Standards and expectancies: contrast and assimilation in judgments of self and others. *Psychology Press* (2005).

[13] E. Aronson, T.D. Wilson, and R.M. Akert. *Social Psychology* Pearson Education; Upper Saddle River, NJ (2003).

[14] H. Heckhausen. Achievement motivation and its constructs: A cognitive model. *Motivation and Emotion* 1 (4). 283-329 (1977).

[15] J. Maddux. Expectancies and the Social-Cognitive Perspective: Basic Principles, Processes, and Variables. In: J. Kirsch, I. (ed.) *How expectancies shape experience*. American Psychological Association. pp.17-39 (1999).

[16] D.O. Sears, L.A. Peplau, J.L. Freedman and S.E. Taylor. *Social Psychology*. Prentice Hall: Englewood Cliffs, NJ (1988).

[17] G. Bierbrauer. *Sozialpsychologie*. Kohlhammer, Stuttgart (2005).

[18] M.R. Endsley, B. Bolté, and D.G. Jones. *Designing for Situation Awareness. An Approach to User-Centered Design*. Taylor & Francis: London and New York (2003).

[19] K. Fischer. What is a situation? *Gothenburg Papers in Computational Linguistics*. 00-05:85-92 (2000).

[20] 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) (2008).

[21] M. Lohse, M. Hanheide, K. Rohlfing, and G. Sagerer. Systemic Interaction Analysis (SInA) in HRI. *Proceedings of Human-robot interaction conference* (2009).