Feedback Guidelines for Multimodal Human-Robot Interaction: How Should a Robot Give Feedback When Asking for Directions?

Nicole Mirnig\textsuperscript{1}, Barbara Gonsior\textsuperscript{2}, Stefan Sosnowski\textsuperscript{2}, Christian Landsiedel\textsuperscript{2}
Dirk Wollherr\textsuperscript{2}, Astrid Weiss\textsuperscript{1}, and Manfred Tscheligi\textsuperscript{1}

Abstract—It is the aim of our research to explore how multimodal feedback can help a robot to carry out itinerary requests effectively and satisfactorily for a human interaction partner. We conducted two studies to evaluate the feedback setup of the Interactive Urban Robot (IURO), which navigates through public space autonomously and finds its way by asking pedestrians for directions. In a Wizard-of-Oz (WOz) experiment with novice users, different feedback modalities and various combinations of them were tested against each other to ascertain the ideal setup of the robot. Subsequently, a cognitive walkthrough with HRI experts was performed to validate the results from the experiment. The results from both studies show that for itinerary requests verbal feedback is most prominent but other feedback modalities may support the conversation by providing reassurance or positive emotions.

I. INTRODUCTION

Imagine, a life-sized humanoid robot is autonomously navigating through public space and proactively approaching pedestrians to retrieve directions. How would we be able to communicate with it and understand its intentions? We conducted two controlled studies to investigate the best feedback configuration for the IURO robot\textsuperscript{1} which has to navigate through a city without prior map knowledge. Thus, all direction information must be retrieved through human-robot interaction from passers-by. To retrieve directions from pedestrians, the robot must proactively initiate and conduct a conversation, which involves providing adequate feedback. In this interaction setup, multi-modal feedback is an important additional cue to achieve human-like interaction. As our studies are aimed at researching feedback as the element that conveys the internal system status of the robot to the human, this paper investigates, which feedback modalities are rated best by human interaction partners.

Autonomous navigation is still a very demanding challenge for mobile robots and particularly in unknown environments a robot will always face knowledge gaps as not everything can be pre-programmed. The ability to detect gaps in its own knowledge and retrieve missing information from human agents is a highly desirable but not yet implemented feature of contemporary robots.

In the context of assistive robots, natural language was already found promising for human-robot interaction \cite{1}. However, given the fact that communication between humans and robots is still far from the instinctiveness of interpersonal communication, the question remains how multimodal feedback can be designed to enable robots to effectively talk to human interaction partners. It has already been shown that feedback is a crucial factor for successful human-robot communication (HRC) in the context of asking for directions. Adequately timed and appropriately deployed feedback was found to foster a vivid and natural flow of conversation \cite{2}, but it is still open how different modalities can be combined best. Independent of the entire scenario, we were interested in feedback in short direction request dialogs, without the dynamics of a real-world setting. The underlying research is aimed at testing combinations of highly different feedback modalities against each other to search for the best configuration: A user study was performed to locate general tendencies and then supplemented with an expert evaluation to deepen the understanding of the results. The findings from both studies were merged to formulate recommendations for how a robot which is asking for directions should provide feedback to a human interaction partner.

II. RELATED WORK

Much state-of-the-art research on feedback in HRI is dedicated to investigate single feedback modalities instead of multimodality. Riek et al. \cite{3} showed e.g. that the type of a gesture affects the reaction time of a human interaction partner (a “take” gestures evokes faster reaction than “follow” or “shake hands”) and that abrupt gestures result in faster reaction than smooth ones. A robot that produces speech and gestures was rated more likable, friendly, and active than a robot that only speaks \cite{4}; but the gestural feedback was also varied, whereas verbal feedback was more or less assumed as the basis. As many scenarios are based on verbal communication, meaning that the human and the robot first and foremost speak to each other, spoken language must be there to achieve communication at all. The interaction gets multimodal by adding a modality, but most studies focus on this additional modality instead of investigating effects that arise from the combination of different modalities.

The impact of non-verbal communication on what is being said has been widely recognized and several aspects have been explored: Non-verbal social cues were found to...
improve the effectiveness in human-robot teamwork and in particular implicit cues resulted in a higher robustness to errors [5]. A robot that can nod and tilt its face was perceived more natural [6]. The researchers also found that a robot that lacks a movable mouth can compensate the missing lip movement by lifting its head slightly. Huang and Mutlu [7] proposed a behavior toolkit to enable robots effective production of social behavior. In their experiments the researchers focused in particular on gaze behavior. They found out that participants have a better recall on a story told by a robot and show better task performance when the robot displays human-like gaze behavior than when it shows delayed, incongruent, or no gaze behavior at all. Lee et al. [8] had a service robot deliver snacks in an office environment and show personalized behavior (remembers people it has interacted with before or what someone ordered previously). They found that people show more social behavior (e.g. flatter the robot or give it a gift), more cooperation, and higher engagement towards a robot with personalized dialog than compared to an unpersonalized robot.

Despite this extensive research interest in feedback, it remains open so far how multiple additional feedback modalities might improve an interaction in a short-term context with non-recurring and non-familiar interaction partners. Our research aims at improving instances of such interaction, when there is no time for mutual adaption and joint grounding takes only place rudimentally.

III. USER STUDY

The overall aim of this WOz experiment was to research feedback in HRC as the element that conveys the internal system status of a robot to a human. Different feedback modalities were tested to find out: (a) which modalities can be considered as “basic” (what must be there to ensure understanding at all), (b) a ranking of the single modalities according to their effectiveness, efficiency, user satisfaction and the user’s cognitive load, and (c) which combinations of modalities, in the sense of multimodal interaction, are most effective, efficient, satisfying and/or have the lowest cognitive load for the user in direction request situations.

A. Experimental Setup

The WOz experiment was set up within subjects, with a total of 20 novice participants. The robotic head EDDIE [9] was used as an interaction partner for the study participants. EDDIE was placed on a table at which the participants were seated, facing the robot (see fig. 1). Directly underneath the robot head an 8” color screen was mounted to display additional information. A pointing device (subsequently referred to as pointer) was mounted above the robot head to indicate directions. The interactions were recorded on video to be used in the follow-up expert evaluation.

1) Feedback Modalities and Tested Conditions: The following feedback modalities were tested to find out which ones are most effective, efficient and satisfying, with the lowest cognitive load for the user in direction-request situations: (1) EDDIE’s verbal utterances (no experimental condition: the dialog was scripted and followed the same structure for all conditions), (2) facial expressions of EDDIE (no experimental condition: the participant’s facial expressions were mirrored through the robot), (3) the pointer to indicate directions, and (4) the screen to display the route graph the robot developed during the conversation. The verbal utterances and facial expressions were not varied to keep the complexity of the experiment on a feasible level, but were set up according to a previous study [10]. When pointer and screen were to be used was strictly predefined.

The design decision of having a pointer instead of arms is grounded in the assumption that arms would implicit more abilities than the robot actually has (e.g. grasping). In its idle state, the pointer performs small random movements at irregular intervals to ensure that the user’s attention is not merely attracted by switching the pointer on and off. The screen is used during a confirmation phase to show the participants on a map what the robot understood. In its idle state, the screen shows a neutral image. To study the quality of task-related feedback, it was furthermore explored if experimentally induced misunderstanding affects the interaction; conditions with misunderstandings are marked with “b”.

In b-conditions, the robot pretended that it misunderstood one item (e.g. the user says “left” but the robot repeats “right”). Table I gives an overview on the conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Pointer</th>
<th>Screen</th>
<th>Misunderstanding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>Idle</td>
<td>Idle</td>
<td>-</td>
</tr>
<tr>
<td>1b</td>
<td>Idle</td>
<td>Idle</td>
<td>-</td>
</tr>
<tr>
<td>2a</td>
<td>On</td>
<td>Idle</td>
<td>b</td>
</tr>
<tr>
<td>2b</td>
<td>On</td>
<td>Idle</td>
<td>-</td>
</tr>
<tr>
<td>3a</td>
<td>Idle</td>
<td>On</td>
<td>-</td>
</tr>
<tr>
<td>3b</td>
<td>Idle</td>
<td>On</td>
<td>-</td>
</tr>
<tr>
<td>4a</td>
<td>On</td>
<td>On</td>
<td>-</td>
</tr>
<tr>
<td>4b</td>
<td>On</td>
<td>On</td>
<td>-</td>
</tr>
</tbody>
</table>

The conditions were tested in pairs (corresponding a- and b-conditions were tested successively) and the participants were only asked questions after each pair of runs. The condition sequences were counterbalanced to prevent learning
effects. To ensure comparable conditions, a wizarding tool was used that strictly followed a predefined script.

2) Study Environment and Equipment: The experimental system architecture consisted of several sub-systems with interconnections realized using RTDB [11] and ROS\(^2\) middleware. The actions of the system were generated by two modules, one for recognition and synthesis of facial expressions according to the Social Motivation Model [12], and one for generating the flow of dialog acts. The latter handles the occurrence of events with timing related to the progression of the dialog including verbal utterances, pointer movements and route graphs presented on the screen. The verbal utterances followed the divided dialog strategy [12], in which the robot asks the participant to provide the directions divided into small segments and verifies them in a confirmation phase. Since a robot with a female voice was perceived slightly better in a previous experiment on HRC in public space [2], a female synthesized voice was used.

3) Research Question and Measures: The subsequently reported studies were performed to answer the following research question: What are the most effective, efficient, and satisfying feedback modalities with the lowest cognitive load for users in direction requests in HRC? The following measures were applied: (1) performance measures - How long does it take to complete the task? (Efficiency); How many misunderstandings of the robot are the participants able to detect? (Effectiveness); (2) physiological measure - Where do the participants look at during the conversation and for how long? (Efficiency); (3) subjective measures - ratings based on the NASA task load index (TLX) directly after each set of two runs. (Cognitive Load); self-defined questions on user satisfaction after each set of two runs and interview with open-ended questions plus Likert-scaled questions at the end of the test (User Satisfaction).

B. Study Procedure

The participants played the so-called “taxi-driver game” with the robot, based on directions that were given to them in the form of a map with an indicated route and destination. As we wanted the participants to fully concentrate on the robot, the map was provided and updated on a separate screen. The map was divided into three segments that were presented on a tablet PC one after the other (see fig. 2), each of which was visible for a short period of eight seconds. The participants were told beforehand that the map segments would disappear and they were asked to memorize them as good as possible.

Each conversation followed the same set of interaction sequences [2] and the same pattern:
(1) The robot said “hello”, introduced itself and asked if the participant was able to explain the directions.
(2) As soon as the participant agreed to help the robot, it asked for the directions of the first map segment.
(3) By touching the tablet PC, the participant could display the first map segment, which was shown for eight seconds.
(4) Then, the participant explained the first segment to the robot. If the robot did not understand the directions it asked the participant to repeat the information. If the robot had understood what the participant explained, it asked if it had arrived at the final destination. Since two other map segments were to follow, the participant answered “no”.
(5) Next, the robot asked for the second segment. Upon touching the tablet PC, the first and the second segment were shown in one picture. The procedure was repeated for the third segment.
(6) After the participant approved that the directions were complete, the robot started to confirm what it understood. During this phase the robot provided feedback according to the condition of the particular run.

C. Results

The 20 participants were counterbalanced regarding gender, their mean age was 27 years (SD 5.85). In the following, the results from the performance, physiological and subjective measures are described.

1) Performance Measures: Table II shows the mean task duration for runs without induced misunderstanding (Kolmogorov-Smirnov test was not significant - normal distribution may be assumed). According to the descriptive results, the dialogs in the condition without additional feedback are less efficient compared to those conditions, in which the robot provides additional feedback with either the pointer, the screen, or both. A trend regarding within-subjects contrasts between the condition with no additional feedback and the three conditions in which additional feedback modalities were used can be observed in a repeated measures ANOVA: \(F(1,19) = 4.06, p = .069\). A larger sample might lead to significant results (Efficiency). To investigate whether the ratio of responses is different across the groups, the detected misunderstandings were analyzed using Cochran’s Q test.

Fig. 2. Sample map divided into three segments

2Robot Operating System, www.ros.org

Fig. 3. Participant explaining the directions to EDDIE

535
Two participants had to be excluded from the analysis as they confused left with right and vice versa. We found a significant difference between the feedback modalities regarding the detected misunderstandings ($\chi^2 (3) = 9.158, p = .027$). A pairwise comparison revealed that significantly less participants detected the robot’s misunderstandings in condition 4b, where it gave feedback with pointer and screen ($p = .032, \phi = -.389$), as compared to condition 3b in which it only used the screen. In condition 4b only 6 participants detected the misunderstanding, whereas in condition 3b 14 people discovered that the robot did not understand every instruction (Effectiveness).

**TABLE II**

**MEAN TASK DURATION (NO MISUNDERSTANDING)**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Feedback Modalities</th>
<th>Mean Duration</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>verbal + facial</td>
<td>0:03:29</td>
<td>0:00:58.79</td>
</tr>
<tr>
<td>2a</td>
<td>verbal + facial + pointer</td>
<td>0:03:11</td>
<td>0:00:49.63</td>
</tr>
<tr>
<td>3a</td>
<td>verbal + facial + screen</td>
<td>0:03:11</td>
<td>0:00:39.11</td>
</tr>
<tr>
<td>4a</td>
<td>verbal + facial + pointer + screen</td>
<td>0:03:10</td>
<td>0:00:42.52</td>
</tr>
</tbody>
</table>

2) **Physiological Measures:** During the participants’ interaction with EDDIE, eye-tracking data was collected to find out where the participants look when talking to the robot. Three areas of interest (AOI) were defined: the robot’s face, pointer, and screen. Table III gives an overview on how much of the time the participants looked at which area (Efficiency).

**TABLE III**

**WHERE DID THE PARTICIPANTS LOOK WHILE TALKING TO EDDIE?**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Face</th>
<th>Pointer</th>
<th>Screen</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>68%</td>
<td>4%</td>
<td>3%</td>
<td>23%</td>
</tr>
<tr>
<td>1b</td>
<td>69%</td>
<td>3%</td>
<td>2%</td>
<td>26%</td>
</tr>
<tr>
<td>2a</td>
<td>70%</td>
<td>3%</td>
<td>2%</td>
<td>23%</td>
</tr>
<tr>
<td>2b</td>
<td>68%</td>
<td>3%</td>
<td>2%</td>
<td>27%</td>
</tr>
<tr>
<td>3a</td>
<td>43%</td>
<td>3%</td>
<td>11%</td>
<td>43%</td>
</tr>
<tr>
<td>3b</td>
<td>64%</td>
<td>2%</td>
<td>14%</td>
<td>20%</td>
</tr>
<tr>
<td>4a</td>
<td>61%</td>
<td>5%</td>
<td>10%</td>
<td>24%</td>
</tr>
<tr>
<td>4b</td>
<td>58%</td>
<td>3%</td>
<td>16%</td>
<td>23%</td>
</tr>
</tbody>
</table>

The face is by far the area that was most often looked at during the interaction, whereas the pointer was hardly looked at, as was the screen in conditions where it was turned off. However, in conditions where the screen was used, an increased interest in the particular AOI can be recognized.

3) **Subjective Measures:** To test if there is a downside to our findings regarding task efficiency, where we found out that additional modalities improve an interaction, we measured the participants’ cognitive load by asking them to complete six rating scales based on the NASA-TLX. A repeated-measures ANOVA, however, showed no significant differences between the conditions (Cognitive Load). The participants rated the scales in general very low. The means over all conditions are (0 = no demand to 100 = maximum demand): mental demand 28.13 (SD 19.72), physical demand 15.31 (SD 15.11), temporal demand 26.56 (SD 20.43), performance 77.50 (SD 21.20), effort 29.13 (SD 18.62), and frustration 25.50 (SD 19.03).

After every set of two runs, the participants were asked which kind of feedback they had received. Whereas all participants noted the verbal feedback, the facial expressions were mentioned by about half of the users. Pointer and screen were hardly mentioned, but the pointer was in a few cases also mentioned in conditions in which it was in idle mode and not indicating directions.

After all eight runs, the participants were asked to rank the four feedback modalities according to their importance (see table IV). Apart from four participants who preferred the screen over verbal feedback, 16 participants stated that they favor verbal utterances as feedback modality. All modalities were ranked second place by some participants, with a slight tendency towards the screen being second most useful. (Despite the fact that the screen was not mentioned by many participants after the actual runs.) Feedback via facial expressions was ranked in third place by most of the participants. 15 participants agreed on the pointer being the least useful feedback modality (User Satisfaction). A Kruskall-Wallis test showed significant differences in the participants’ ranking ($\chi^2(3, N=20) = 45.27, p = .00$). Follow-up tests were conducted to evaluate pairwise differences among the four modalities, controlling for Type I error across tests by using the Holm’s sequential Bonferroni approach. The pairwise comparisons showed that all modalities significantly differ from each other in terms of their ranking, except for facial expressions and the screen. It can thus be said, that verbal feedback is perceived as the most important feedback modality, whereas the pointer is ranked least important. Facial expressions and screen are ranked in between, but it can not clearly be stated which one is more important.

**TABLE IV**

**PARTICIPANTS’ RANKING OF FEEDBACK MODALITIES**

<table>
<thead>
<tr>
<th>Modality</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal</td>
<td>16</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Facial Expression</td>
<td>0</td>
<td>5</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>Pointer</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>Screen</td>
<td>4</td>
<td>8</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>

4) **Summary of Results:** With our WOz user study we wanted to research how a speaking robot that asks a human for directions could be improved by equipping it with additional feedback modalities. It was the aim to find out which modalities are most effective, efficient, and satisfying, with the lowest cognitive load for a human interaction partner. We could show that adding feedback modalities is beneficial especially in terms of reassuring and comforting the user. Too many modalities, however, outdo each other and are thus counterproductive. A trend could be observed regarding efficiency that if the robot provides feedback via additional
modalities, the user is able to provide the requested directions faster. However, no difference could be made out between the pointer, the screen, or the combination of both. The face is the area of the robot where the participants are most likely to look at during an interaction. The pointer on the contrary is hardly looked at, same as the screen when turned off. In conditions where the screen was used, people looked at it more frequently. As to effectiveness, it could be shown that significantly less participants detected the robot’s misunderstandings when it provided feedback via all modalities. Regarding the participants’ cognitive load, no significant differences could be detected between the modalities. This might be due to the small sample size or due to single feedback modalities being maybe too implicit. Finally, inquiring on the user satisfaction revealed that the participants favor verbal feedback over the screen. Feedback via facial expressions and via the pointer were ranked in third and fourth place. The experiment shows that for direction requests verbal feedback is most important, but it can be ideally supplemented by providing additional feedback via a screen to reassure the participants and via facial expressions to make the conversation more pleasant.

IV. EXPERT EVALUATION

To deepen and verify the insights we gained during the user study, we decided for a follow-up cognitive walkthrough, in which at least two experts try to imagine how a typical user would complete a certain task and in doing so assess the usability of a system. It was adapted for HRI by Weiss et al. [13]. Upon going through the task, the experts ask four questions and as soon as one of them is answered with “no”, a usability problem is detected. The experts were asked to watch video sequences of the original interaction and put themselves in the participant’s position. Three HRI experts performed our cognitive walkthrough, which was moderated by one researcher.

A. Study Design

The cognitive walkthrough was based on video data that was recorded during the WOz experiment. A video sequence of one run from each of the four feedback conditions was shown to the experts, for each of which a separate walkthrough was performed. One video sequence showing an instance of misunderstanding was also shown to the experts.

The robot interacted according to four interaction sequences - introduction, giving/receiving directions, confirmation, and conclusion [2]. At the beginning, the experts were familiarized with the sequences, the overall IURO scenario, and the initial WOz experiment.

B. Study Procedure

Each video was divided into the four interaction sequences. Every sequence was then analyzed as follows:

- Are the robot’s intentions clear?
- Is it visible what the robot is doing at all times?
- Did the robot understand what the user said?
- Will the robot be able to find the destination? Did the feedback setup influence the conversation and how?

C. Study Results

The HRI experts identified 22 problems: 4 in the introduction phase, 2 during giving/retrieving directions, 13 in the confirmation phase and 3 during the conclusion. Table V shows how the experts assessed the feedback combinations in the confirmation phase, which was when the robot provided feedback depending on the respective condition.

| TABLE V |
| Feedback Assessment in the Confirmation Phase |
| Feedback Condition | No. | Problem Description |
| verbal + facial | 5 | repetition divided into segments is unnatural, dialog structure too strict, correctness inquiry after every segment is too much, not comprehensible if the robot understood the directions, more natural verbal feedback |
| verbal + facial + pointer | 4 | pointer position is inappropriate, pointer movement is delayed and unclear, mimics distract from the pointer |
| verbal + facial + screen | 2 | map is unclear, map is only shown during the repetition |
| verbal + facial + pointer + screen | 2 | four modalities might be too much, robot refers to the “destination” |

The experts provided valuable suggestions for optimization, which are of course in part specific to the tested setup and cannot be generalized as such: The verbal feedback could be improved by making the robot react faster and by a less strict dialog structure that allows for more variation. The screen could be improved by dynamically drawing the map when the robot speaks. It should be more abstract to bring across that the drawing is what the robot understood (and thus could be wrong if the robot mistakes something). The facial expressions could be improved by making the robot smile more often or even make it smile as a positive feedback when it understands the directions. The randomized pointer movement is considered as too distracting and not interpretable. The pointer should furthermore be synchronized with the verbal feedback. The experts suggested to position the pointer in an area where gestures are expected (below eye level). Two remarks were made regardless of the single modalities: (1) The robot should in any case provide intermediate feedback to show the interaction partner e.g. that it is currently processing information; (2) using four feedback modalities at a time might be too much as e.g. mimics may distract from the pointer.

Finally, the experts were asked to rank the four feedback modalities according to their importance. Their ranking was similar to that of the users in the experiment: (1) verbal, (2) screen, (3) facial expressions, and (4) pointer.

V. FEEDBACK RECOMMENDATIONS

For short-term interaction in public space such as itinerary requests we can infer the following from our study results: Both, the user study and the expert evaluation showed that a robot which provides additional feedback to supplement its verbal utterances is perceived positively. However, several feedback modalities operating in parallel are not per se
favorable, but require good system design and alignment to ensure their supportiveness of the interaction. The following recommendations for feedback of a robot in a speech-based scenario in public space can be given:

**Provide feedback in areas where it is expected:** The face is by far the area that most participants look at during most of the conversation; the pointer is hardly looked at whereas the screen draws some attention and functions as a backup channel to provide reassurance.

**Make verbal feedback the most important modality:** For itinerary requests, verbal feedback is the most important component to make such dialogs work. However, additional feedback modalities provide the participants with increased comfort and reassurance. Facial expressions coordinated with the dialog are likely to increase empathy, whereas providing the depiction of what the robot understood on a screen helps the participants to detect if the robot understood them.

**Provide additional feedback through alternative modalities, but make sure that they are readable:** It might be reasonable to use a pointer for gestures. However, this modality requires extensive research to ensure it is interpretable, e.g. in terms of positioning and movement manner.

The results of the studies as reported above are liable to the following limitations: First, a larger sample could result in more significant results. Second, testing direction retrieval dialogs in a lab setting is suboptimal as such dialogs are clearly oriented towards a real world setting which is highly dynamic, where actual grounding processes take place and gestures may be used to reference to a certain point in the landscape and not just to refer to a theoretical route.

**VI. CONCLUSION**

This paper reported on two conducted studies in which the human-like feedback modalities “verbal utterances” and “facial expressions” were combined with two non human-like modalities to explore their eligibility for proactive information retrieval from task-related HRI.

A trend could be detected that interactions with only human-like modalities are less efficient regarding task completion time than if non human-like modalities are used in addition. With this results we do not want to say that human-like modalities are less efficient but rather that empathy-like cues such as facial expressions in general do not account for raising efficiency. However, these results have to be verified regarding their significance by a larger sample. The total ranking of all modalities reveals clear preferences for the verbal utterances. This, again, confirms natural language to be the modality of choice for information retrieval in this interaction domain. Although non human-like, displaying the retrieved route knowledge of the robot on a screen turns out to be a valuable means of reassurance for users, and thus is rated even better than the modality of human-like facial expressions. Hence, the screen seems to provide a “bridge of reassurance” between the verbal utterances which transport the exchanged information between a human and a robot, and facial expressions which create the motivating social background for this process. The pointer as additional

---

**REFERENCES**