Exploring Users' Reactions Towards Tangible Implicit Probes for Measuring Human-Robot Engagement

Mohammad Obaid[1], Yuan Gao[1],
Wolmet Barendregt[2], and Ginevra Castellano[1]

[1]Department of Information Technology, Uppsala University, Uppsala, Sweden
[2]Department of Applied IT, Gothenburg University, Gothenburg, Sweden

Abstract. In this paper, we present an exploratory study of the use of tangible implicit probes to gauge the user's social engagement with a robot. Our results show that users' paying attention to the robot's implicit probes is related to higher social engagement, but also that introducing implicit probes can lead to a more positive interaction with a robot. As we observed that users in our study started paying more attention to the implicit probes after they had encountered them, the need for careful design to capture changes in social engagement over time is justified here. Finally, we discuss some of the user recommendations to design better implicit probes.

Keywords: Robot; social engagement; task engagement; implicit probes.

1   Introduction

Social robots are increasingly being used in several interactive domains where the quality of the social engagement between the robot and its user/s, such as educational robots and robotic companions is key. An important characteristic in a robot is thus being able to assess whether the user is socially engaged with it, and adapt to it accordingly [1].

Traditionally, research efforts have been directed mainly towards using sensor technologies to automatically or semi-automatically detect users' affect, engagement and attention towards robotic agents, such as the work by Shangvi et al. [10] and Salam et al. [9]. One of the drawbacks of using sensor technologies is that they are prone to error and unreliable when data is noisy. Additionally, users may perceive wearing a sensor as too intrusive. Although it might be possible to directly ask the users how engaged they are with the robot, this may be too interrupting. Rather, one would like to establish ways for the robots to assess the user's engagement with them more implicitly. There are very few attempts by researchers to investigate how a robot can assess its user's engagement in an implicit way that is also context-specific, such as [4,5]. Most of these attempts are based on dialogue-based "implicit probes", which require the user to attend to specific probes in a form of a dialogue or questions. In this paper, we aim

to explore a novel way of eliciting information about users' engagement with the robot by using tangible implicit probes. In our proposal, information is also provided directly by the users without using sensor technologies, but rather than being dialogue-based it is based on tangible actions with the robot.

2   Related work

Definitions of engagement are many in Human-Computer and Human-Robot Interaction (HRI) [5]. While engagement is frequently operationalized by means of measures of visual attention, it is important to distinguish conceptually between engagement and attention. Engagement is a complex phenomenon, a construct consisting of both cognitive (attention, concentration) and affective components (enjoyment) [6][1]. In this research, we are interested in social engagement in HRI, where social engagement is defined by Rich et al. as "the value that a participant in an interaction attributes to the goal of being together with the other participant(s) and continuing the interaction" [8].

Peters et al. [7] have explored the fundamental factors behind the concept of social engagement in order to model engagement capabilities for machines. The fundamental factors they describe are the phases of engagement, the focus of engagement, and interest. In addition, they have aimed to relate different views on engagement to the perception-cognition-action loop. Approaches for measuring engagement in HRI rely primarily on machine-learning based models using data from sensors, conveying information about behavioural signals such as facial expressions, eye gaze, and body posture [5]. A disadvantage of these approaches is that they are prone to suffer from poor performance. This is due to issues such as noisy sensors and training data, limited sensitivity to individual differences in expressing affect (for example, to different personalities or cultures), models that fail to capture relationships between different features and to account for the context of the interaction.

Previous research has proposed the concept of engagement probes [3][5]. Such probes form a non-intrusive and embedded method to collect informative data at different HRI stages. The idea is to embed mechanisms in the HRI cycle by including interaction events that allow robots to proactively collect data about the user's engagement. Such events initiated by the robot may act as implicit queries aiming to assess, for example, whether the user diverts gaze from the task towards to the robot, maintains the gaze towards the task but adapts their actions to follow the robot's guidance, or ignores the robot's guidance completely. These queries may help elicit responses that are known to be associated with low or high levels of engagement [12], and to understand if the user is engaged with a given task, or with both the task and the robot. In this context, task engagement is defined as a combination of the three elements: attention, concentration, and enjoyment[12]. Thus, the aim of the probes is to elicit specific user responses (e.g., verbal or non-verbal behaviours) that are expected, based on results from the social sciences literature, to be highly meaningful for the purpose of discriminating a specific affective state.

Building on the previous work on probes, we aim to use implicit probes to gauge the user's social engagement, assuming that paying attention to the probes signifies this. A first exploration of the usefulness of the implicit probes to measure social engagement is to determine whether this assumption is correct. Our first **RQ1** is therefore: Is there a correlation between the users' attention to the implicit probes and the subjective assessment of their social engagement with the robot? If RQ1 is answered positively, we would like to know the following **RQ2**: Are there certain types of implicit probes that are more accurate than others in order to determine the users' social engagement with the robot?

However, it is possible that the implicit probes themselves influence either the social or the task engagement, and alter the users' assessment of the interaction/robot. For example, the implicit probes could cause the users to be less engaged with the task, or more engaged with the robot. The third research question is thus: **RQ3**: Do the implicit probes affect the users' levels of social/task engagement or impressions of the interaction and impressions of the robot?

Finally, in this study, we are interested in the design of such implicit probes. Although RQ2 results may indicate which kinds of probes are most useful to differentiate the user's social engagement, we also want to know **RQ4**: How do people react to the different kinds of implicit probes and what are their subjective impressions of them?

3 User study

To answer the research questions, we conducted a between-subject study with two conditions (with/without implicit probes) with 26 participants. The main task for all users in this study was to assemble a LEGO set using an incomplete instructional manual booklet, together with a robot that was controlled by a Wizard-of-Oz method. The robot's role was to support the participants by providing them with the missing instructional manual pages to allow them to accomplish the task. We chose to use the Robot Bird model from the LEGO Creator Robo Explorer Building Kit, which has 37 building instructions (15 pages). We omitted 5 instructional pages and replaced them with blank pages. We then formatted these pages to be displayed on the robot's screen. The wizard observed the participant's progress through the instructional manual and executed a command to displayed the missing page on the robot's screen when a missing page was flipped (blank page). Overall, the WoZ executions were consistent in the way the commands were activated in relation to the wizard's observations.

Condition A: In condition A we used implicit probes that invited the user to interact with the robot's body just before or after a missing page. In total we had four implicit probes; two of them were based on gestural interactions with the robot: a high five gesture and a power hand gesture. We chose the two gestural interactions to illustrate the robot's support on the user's achievements. The other two probes were touch-screen based, where the user pushed a button on the robot's screen in order to charge its battery. We chose the touch-screen interactions to see if the user cares about the robot when an alert is activated.

In this context, the first touch-screen probe was accompanied with a beep sound while the other used a spoken dialogue to alert the user (for example, "I want to feel more energetic"). Figure 1 illustrates the four implicit probes used in our study. In this condition, the wizard executed a command to activate an implicit probe event, which only started 20s after the command was executed to make sure the user had started working with the flipped page.
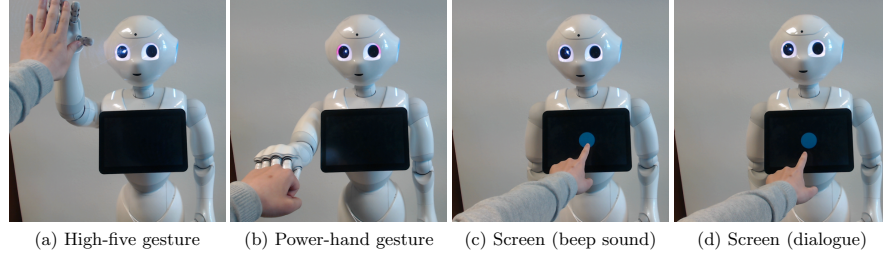

(a) High-five gesture   (b) Power-hand gesture   (c) Screen (beep sound)   (d) Screen (dialogue)

Fig. 1: The four implicit probes used in the presented exploratory study

Condition B: In condition B, all the probes presented in condition A were replaced with dialogue statements that were of an informative nature such as "I think you are doing the 7th page, you are almost half way". The wizard followed the same procedure as for condition A but activated the informative statements instead of the implicit probe events. To keep the two conditions consistent we designed the process so that the execution of the events was dependent on the page being flipped. Figure 2 illustrates the whole process and the different events that the wizard was responsible for activating after each page.
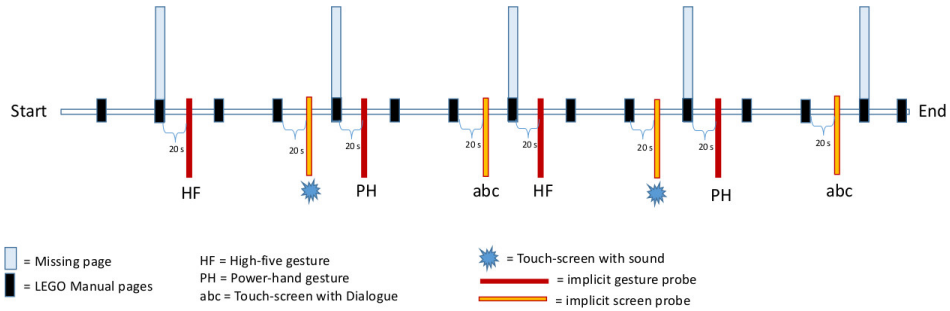

Fig. 2: The task's time-line and the events the wizard is responsible for executing in Condition A (for Condition B the events are replaced by informative statements)

Participants: In total, 26 participants completed the study. Condition A had 14 participants (6 females) with an average age of 30.0 years (SD=6.8), and condition B had 12 participants (9 females) with an average age of 25.7 years (SD=3.2). Participants were mostly university students with IT backgrounds.

Additionally, we asked participants to rate their experience with playing with LEGO and with interacting with a robot on a 5-point Likert scale. There were no significant differences between the conditions related to previous LEGO experience or robot experience (p> 0.05).

Apparatus: The study was setup in a 3x4m room, and was equipped with two cameras, the robot Pepper, a work desk, and a LEGO set. The room was split into two with an office partition wall, where one part was used to run the study and the other was for the wizard. The robot used in our study is the Pepper robot from Softbank Robotics, which has a height of 120cm. The robot's events and animations were programmed using the Choreographe software. The robot's pre-defined autonomous-life option was enabled to give it a lively idle feeling, while all other actions were controlled and activated by the wizard. For the predefined wizard actions, we used Python and an SDK called naoqi provided by Softbank robotics.

Procedure and measurements: each session started by the study supervisor explaining the setup and asking the participant to read and sign a consent form. If signed, the participant was asked to fill out a demographic questionnaire before explaining the task they had to do with Pepper. The participant was given a brief introduction on what would be their task and how the robot would help them throughout the session. For condition A, the participant was given an introduction on how they could interact with the robot when the four probes (described in Figure 1) were executed. Upon completing the task, the participant was asked to fill out the study questionnaire.

To answer the first two research questions we collected data on the participants' response to each of the implicit probes. Each probe was recorded as either 1 or 0 indicating whether the user responded to that probe. Since there were eight implicit probes in each session, the user thus received a score between 0 and 8 on how much overall attention was given to the probes. After the study session, we then asked the users to assess their own social engagement using a 7-point Likert scale questionnaire that has 10 items. To answer the third research question, we also assessed the users' task engagement on a 7-point Likert-scale questionnaire that has 7 items. Both engagement questionnaires were adopted with adjustments from [4][11] (see Table 1).

We also adopted two sub-questionnaires from Clabaugh and Matarić [2] to determine the users' state of mind and their impression of the robot. The first questionnaire contained 10 terms for the users to rate on a 5-point Likert scale: 'Relaxed', 'Stimulated', 'Calm', 'Engaged', 'Excited', 'Comfortable', 'Intelligent', 'Optimistic', 'Annoyed', and 'Frustrated'. While the second sub-questionnaire asked users to rate, on a 5-point Likert scale, their impression of the robot Pepper on the following terms: 'Likeable', 'Intelligent', 'Life-like', 'Interactive',

Table 1: The social and task engagement questionnaire items

	Social Engagement
1	I would like to play another game with Pepper
2	I feel Pepper was like a partner in the task
3	I believe Pepper cared about whether I did well in the task, or not
4	I was worried about Pepper when its energy started running low
5	I would have felt badly if I was unable to energize Pepper
6	I wanted to keep helping Pepper
7	I felt like Pepper and I were part of the same team
8	I found Pepper considerate
9	I found Pepper helpful
10	I found Pepper friendly
	Task Engagement
1	I enjoyed this task
2	I found this task challenging
3	I would have liked to continue doing this task
4	It was important to me to do well on this task
5	I found this task easy to understand
6	I found this task interesting
7	I have had previous experience with a similar task

and 'Social'. In addition, participants had the option to give written comments at the end of the questionnaire forms.

Finally, to answer the last research question, we conducted a semi-structured interview to gain feedback about the implicit probes used. The interview had questions that focused on the gesture-based implicit probes, screen-based implicit probes, and advantages/disadvantages of the different probes used: (1) 'Did you feel the urge to attend to Peppers actions? Did you believe that Pepper was running out of energy? And why?', (2) 'Did you feel good about the high five and power hand interactions? Where did it feel wrong? And why?', (3) 'What other suggestions do you have to improve these interactions?', (4) 'Did you find Pepper annoying? Did you feel it disturbed/interrupted you?', and (5) 'Did you care about making Pepper happy? Did you care about Pepper in general?'.

4   Results

RQ1 - Correlation between probes and social engagement: As expected, paying attention to the implicit probes was strongly related to subjective social engagement r=.72, p<0.01.

RQ2 - Types of probes related to social engagement: There were eight implicit probes of four different types. Based on the results of the study, only the first high-five and the first interactive touch-screen probe with spoken dialogue were strongly related to subjective social engagement r=.77, p<0.01 and r=.65, p<0.01. Other types of implicit probes, and even subsequent implicit probes of

the same type were not related (p> 0.05) to the participants' social engagement.
RQ3 - Impact of the Implicit probes: Independent samples t-tests were conducted to compare the task engagement, social engagement, the users' ratings of their personal state, and how they felt towards Pepper during the session between the two conditions. The only significant differences between the two conditions were for the terms Relaxed (t=2.23, p<0.05), Calm (t=2.23, p<0.05), and Intelligent (t=2.15 p<0.05), related to the users' own state of mind during the interaction. Figure 3 shows the participants' assessment of their state of mind during the interaction for both conditions.
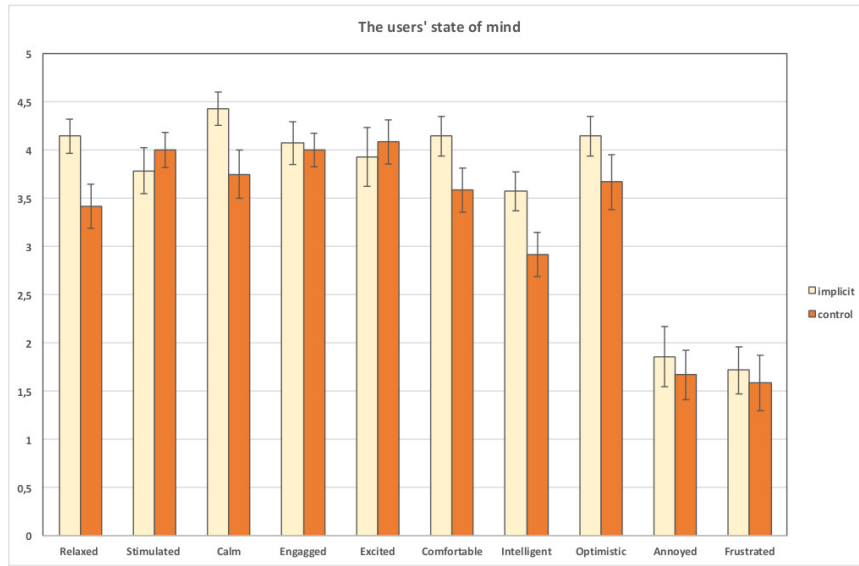

Fig. 3: Users' ratings of their state of mind during the interaction

RQ4 - Reactions to and impressions of different probes: To answer the final question we describe the participants' reactions to the implicit probes as observed by the study facilitator, and as gauged through the explicit questions.

Pepper's screen-based probes: Of the 13 participants who answered this question, five said that they cared about Pepper and that they therefore wanted to make sure that it had enough energy. However, eight participants said they did not care about Pepper but they thought attending to its energy was a task they had to perform. As one participant said: "It felt like a pop-up notification. I did not believe it was running out of energy".

The observations showed that the beep sound event was noted to have a strong impact on attracting the participant's attention to touch the button on Pepper's screen. Participants usually turned towards the robot in a swift way to touch the screen whenever they heard the beep sound. Generally the participants'

facial expressions were rather neutral when attending to the beep sound, and this might explain that the participants really felt that it was more like a pop-up window that they had to close, than feeling care for the robot's energetic state.

Regarding the touch-screen with the spoken dialogue probe, it was noted that participants were more expressive when the robot told them that it was running out of energy, and several tried to initiate small talk regarding the situation. In general, participants seemed more expressive towards the touch-screen with spoken dialogue, which may have been a result of them liking Pepper's voice, as indicated by some of the feedback. Video-based analysis should be carried out to confirm this finding.

Pepper's hand probes: Of the 12 participants who answered the question, five enjoyed the hand interactions with Pepper (high-five and power-hand) and felt they were quite natural. Three did not like the hand interaction, because it was unnatural, or the timing felt slightly off. Finally, four participants felt neutral about them. Two (one who was rather positive and one who was rather negative) were afraid to break the robot's hand. Two participants also indicated that they might prefer fewer bodily interactions, and would rather have the robot talk more and be involved in a conversation. One of the participants also remarked: "Sometimes the power-hand and the high-five felt a bit forced - if Pepper could see if I figured something out, then a high-five would be more appropriate". This suggests that the implicit probes should be more context-sensitive.

The participants seemed to find both gestural probes exciting and more interactive compared to the touch screen probes. In some cases, for example, participants were trying to initiate a high-five gesture by raising their hand and telling Pepper to go for a high-five. However, in a few cases the animation of the gestures seemed a bit slow, for example, some participants tried to follow the power-hand fist as it was raising towards them. As indicated by some of the comments, and also observed, the physical embodied interaction with the robot could be perceived as the most likable way to interact with Pepper when compared with a touch screen interaction or dialogue.

5   Discussion

The results presented above indicate that paying attention to the implicit probes may relate to users' own assessment of their social engagement. However, it also seems that this may only hold for probes given in the beginning of the interaction. Similar probes given later on are not related to this subjective overall assessment. In fact, users paid attention to almost all probes given later on. This could on the one hand mean that they really are more socially engaged during later parts of the session, but it could also mean that they are more prepared for the occurrence of a specific implicit probe, or that they come to expect an implicit probe at regular intervals. This might suggest that implicit probes should vary and come more unexpectedly throughout the interaction. Probably, they should come at times when the robot needs to gauge the user's engagement with it in order to adapt its behavior accordingly. However, the participants only received

eight probes, two sets of four, during the session. Although they paid attention to all probes that were given the second time, it is possible that they would forget to pay attention to all probes during a longer session. As one of the participants indicated: "if this had been a longer task, I would not have attended to the probes". This could possibly lead to a better measurement of the users' changes in engagement, provided that the probes could be given at the right time.

This study has limitations that need to be mentioned. First, this was a small-scale study, involving only a small number of participants, a limited task and a limited number of implicit probes. Second, the study did not address the question whether paying attention to implicit probes relates to in-the-moment changes in social engagement. Third, the study had a fixed set-up, meaning that the probes did not appear in a random order, which might have resulted in a learning effect. A different set-up with more frequent probes would be needed to investigate this. Finally, we assumed that caring for the robot was a sign of social engagement similar to caring for a human. However, we did not investigate in detail whether it was a sign of task engagement instead; some comments from the participants indicated that it actually was. Addressing these limitations in further studies should lead to further insights regarding the use of implicit probes.

6   Conclusion and future work

This exploratory study has investigated whether it is possible to use implicit probes to gauge the user's engagement with a robot while performing a task. The results indicate that this may be possible because there is a strong relation between paying attention to the implicit probes and the users' own assessment of their social engagement. However, this seems to hold only for the probes given at the beginning of the session. Further research is needed to determine whether a longer exposure to the implicit probes allows users to ignore them at times when they are not very socially engaged, which would be necessary to use them as a way to adapt the robot's behavior based on the user's engagement. However, it should also be noted that this could be a sign of high task engagement instead. To validate whether implicit probes could be used to really capture changes in the user's social engagement, a comparison using sensor data should be performed.

While the implicit probes in this study did not seem to have an impact on the users' social or task engagement, or their attitude towards the robot, this needs to investigated in more detail, also taking into account that there may be a trade-off between social engagement and task engagement. Overall, we think this study indicates that the use of implicit probes can be worthy of further investigation. We would suggest looking further into a larger variety of possibly more subtle types of implicit probes, given at irregular times whenever it is crucial for the robot to understand the user's engagement. Determining the right time for a probe is something that is not completely obvious and may well vary from user to user. Possibly, the probes could also be used to gather more information than just the fact that the user is paying attention to them, e.g. as

presented by Clabaugh and Matarić [2] to elicit learning-sensitive information from the user.

Acknowledgments

References

1. Castellano, G., Leite, I., Pereira, A., Martinho, C., Paiva, A., Mcowan, P.W.: Context-sensitive affect recognition for a robotic game companion. ACM Trans. Interact. Intell. Syst. 4(2), 10:1–10:25 (Jun 2014)
2. Clabaugh, C., Matarić, M.J.: Exploring elicitation frequency of learning-sensitive information by a robotic tutor for interactive personalization. In: 2016 25th IEEE Int. Symposium on Robot and Human Interactive Communication (RO-MAN). pp. 968–973 (Aug 2016)
3. Corrigan, L.J., Basedow, C., Küster, D., Kappas, A., Peters, C., Castellano, G.: Mixing implicit and explicit probes: Finding a ground truth for engagement in social human-robot interactions. In: Proc. of the 2014 ACM/IEEE Int. Conference on Human-robot Interaction. pp. 140–141. HRI '14, ACM, New York, NY, USA (2014)
4. Corrigan, L.J., Basedow, C., Küster, D., Kappas, A., Peters, C., Castellano, G.: Perception matters! engagement in task orientated social robotics. In: Proc. of the 24th IEEE Int. Symposium on Robot and Human Interactive Communication (RO-MAN). pp. 375–380. IEEE (2015)
5. Corrigan, L.J., Peters, C., Küster, D., Castellano, G.: Engagement Perception and Generation for Social Robots and Virtual Agents, pp. 29–51. Springer International Publishing, Cham (2016)
6. O'Brien, H.L., Toms, E.G.: What is user engagement? a conceptual framework for defining user engagement with technology. Journal of the Association for Information Science and Technology 59(6), 938–955 (2008)
7. Peters, C., Castellano, G., de Freitas, S.: An exploration of user engagement in hci. In: Proc. of the Int. Workshop on Affective-Aware Virtual Agents and Social Robots. pp. 9:1–9:3. AFFINE '09, ACM, New York, NY, USA (2009)
8. Rich, C., Ponsleur, B., Holroyd, A., Sidner, C.L.: Recognizing engagement in human-robot interaction. In: Proc. of the 5th ACM/IEEE Int. Conference on Human-robot Interaction. pp. 375–382. HRI '10, IEEE Press, Piscataway, NJ, USA (2010)
9. Salam, H., O. Celiktutan, O., Hupont, I., Gunes, H., Chetouani, M.: Fully automatic analysis of engagement and its relationship to personality in human-robot interactions. IEEE Access Journal 5, 705–721 (2017)
10. Sanghvi, J., Castellano, G., Leite, I., Pereira, A., McOwan, P.W., Paiva, A.: Automatic analysis of affective postures and body motion to detect engagement with a game companion. In: Proc. of the 6th Int. Conference on Human-robot Interaction. pp. 305–312. HRI '11, ACM, New York, NY, USA (2011)

11. Serholt, S., Basedow, C.A., Barendregt, W., Obaid, M.: Comparing a humanoid tutor to a human tutor delivering an instructional task to children. In: 2014 IEEE-RAS Int. Conference on Humanoid Robots. pp. 1134–1141 (Nov 2014)
12. Shernoff, D., Csikszentmihalyi, M., Schneider, B., Shernoff, E.: Student engagement in high school classrooms from the perspective of flow theory. School Psychology Quarter 18, 158–176 (2003)