Expressive, Mobile and Interactive Robot for Improving Human-robot Interaction

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

*CHI 2020 Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.*

© 2020 Copyright is held by the owner/author(s).

ACM ISBN 978-1-4503-6819-3/20/04.

DOI: https://doi.org/10.1145/3334480.XXXXXXX

\*update the above block & DOI per your rightsreview confirmation (provided after acceptance)

|  |  |
| --- | --- |
| First Author  University of Author  Authortown, CA 94022, USA  author1@anotherco.edu  Second Author  VP, Authoring  Authorship Holdings, Ltd.  Awdur SA22 8PP, UK author2@author.ac.uk  Third Author  Fourth Author  Lēkhaka Interaction Labs  Bengaluru 560 080, India  author3@anotherco.com author4@hci.anotherco.com | Fifth Author  YetAnotherCo, Inc.  Authorton, BC V6M 22P Canada  author5@yetanotherco.ca  Sixth Author  Université de Auteur-Sud  40222 Auteur, France  author6@author.fr  Seventh Author  Department of Skrywer,  University of Umbhali,  Cape Town, South Africa author7@umbhaliu.ac.za |

# Abstract

In this paper, we explore the emotional behavior of a robot and its effect on human-robot interaction. We introduce the *EMI* platform, an “Expressive, Mobile and Interactive robot”, equipped with a rear-projected expressive face and the ability for voice-interaction. We conduct an experiment with the EMI robot during a public event, in which attendees were given the option to interact with the robot. Following the operation of the robot we asked participants a series of questions to study the effect of the robot’s expressions. From the survey responses, video observations and informal interviews we deduced key design decisions in EMI that help entice users and increase user appreciation towards the robot. These design choices can help the future development of robots operating in human inhabited spaces.

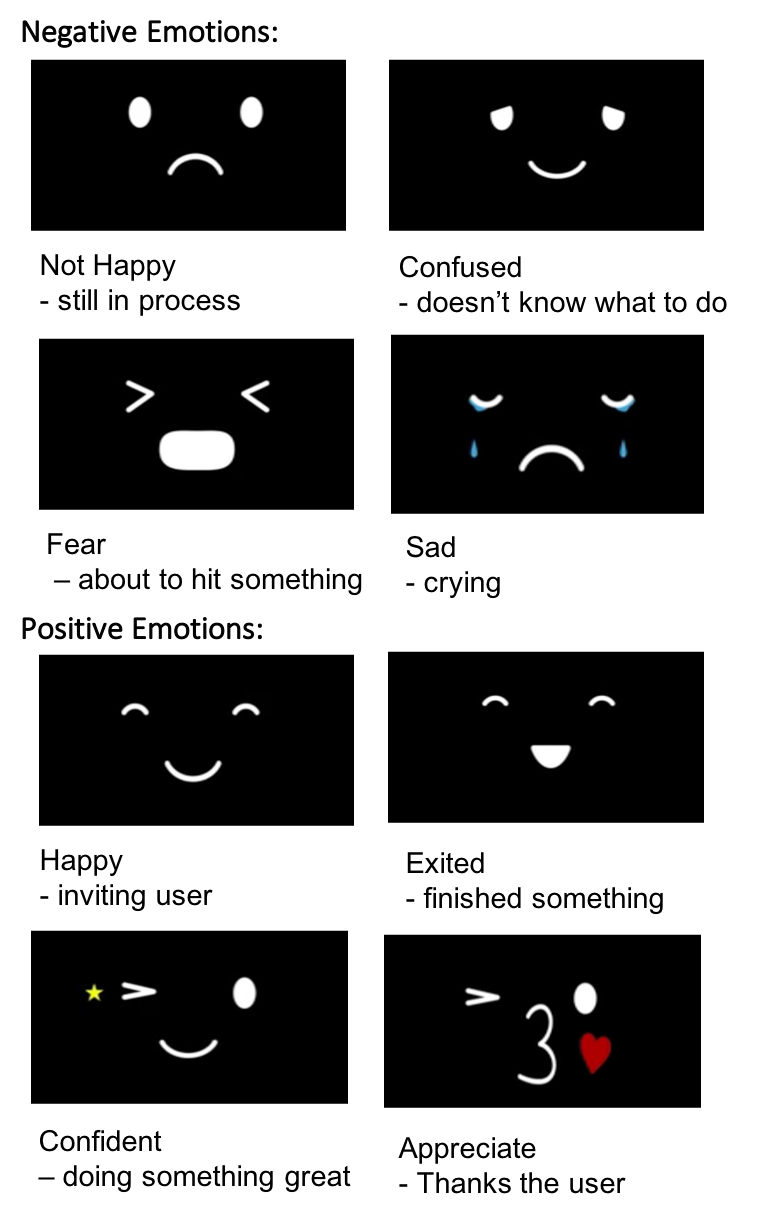
# Author Keywords

Authors’ choice; of terms; separated; by semicolons; include commas, within terms only; required.

# CSS Concepts

• **Human-centered computing~Human computer interaction (HCI)**; *Haptic devices*; User studies; Please use the 2012 Classifiers and see this link to embed them in the text: <https://dl.acm.org/ccs/ccs_flat.cfm>

# Introduction

In human-robot Interaction (HRI), there is growing interest in developing solutions dealing with the social interactions between humans and autonomous robots. As robots operate in human inhabited spaces, the environment can become difficult to interpret and complex to navigate. One approach to handle a situation in which a robot reaches the limits of its capabilities is to enlist human assistance in order to achieve its tasks. For example, a robot may ask for help removing insurmountable obstacles, or working around a self-diagnosed failure. We therefore address the question of how a robot can effectively entice a person to support itself, in situations where there is no designated human operator. At the same time, if the robot is unable to perform the requested tasks in a timely manner, this may lead to user disengagement [Feeling]. We address questions such as: Are there robots that humans are more tolerant of? What features increase human tolerance of a robot? We look at the degree of user tolerance towards the EMI robot was and which design choices have affected this.

# Related Work

Figure []: the booth (with a maze in the middle) and we prepared two EMI robots.

Users will first notice the robot’s appearance, which affects their attitude towards the robot and engagement with it. Amusing, cartoon-like features [A Cross-cultural Study: Effect of Robot Appearance and Task] and a smooth face [Robot’s] have been suggested to increase user preference. It was also shown that expressive communication with can help user engagement in the interaction, especially if the robot produces errors, is slow in carrying out a task, or appearing as being too intelligent [Feeling Committed to a Robot]. Emotional expressions and human-like characteristics help to mitigate such unexpected robot behavior.

Figure []: EMI’s different emotion status and the example usage

A number of technologies has been proposed to convey emotional expressions, including mechanical robotic faces, such as Kismet (shown in figure []), or smooth faces, such as Kasper [] (shown in Figure []). The smooth face also displays animated features using a pixel screen, such as Forpheus [] and rear projected objects. In addition, some robots focus on behavior to convey emotions, such as dancing robots like [keepon].

## rear projected robotic face

Several systems have adapted projecting a robotic face in order to increase expressiveness.

[Mask-bot”] proposed a life-size talking robotic head, “mask-bot” with the ability to express richer facial behaviors than a flat screen. Frederic proposed a robotic face using a retro-projected animated face to study its influence on user ability to interpret its gaze direction. The study did not compare with a flat screen directly, but highlights the importance of gaze direction in design for HRI. Samer al Moubayed developed Furhat, which also used a projected animated head. In their study, the proposed method increased the intelligibility significantly by supporting the visual component of speech perception (lip reading), compared to flat screen.

Prior research focused solely on expressive facial emotion or body motion. This study takes one step further by studying the complete emotional design, robot face, voice and motion, and its effect on human-robot interaction. In this paper, we carry out an experiment in a public event to analyze our proposed solution.

# EMI Robot

|  |  |  |
| --- | --- | --- |
| Environment | Voice Responses | Emotion |
| EMI receives the speech commands from users | “Yes, sure”, “I see!” | smile |
| EMI cannot understand the commands from users | “Sorry, I don’t understand”, “One more time, please” | Sad |
| The received commands will make EMI hit the wall in the maze | “Be careful”  “Watch out” | Fear |
| EMI arrives at key points | “so far so good” | Confidant |
| EMI finish the maze | “Thank you, my friend.” | Appreciate |

We developed a new robot called EMI (Expressive, mobile and interactive). We crafted the facial animation, voice and motive expressiveness to form an identifiable personality.

* Facial expression: EMI is able express different emotions, in conjunction with voice output, shown in Figure [].
* Voice communication: Similar to smart speaker devices, EMI is equipped with wake-up word detection, command recognition and natural language understanding, to enable natural communication.
* Motive expressions: EMI is built using the differential drive Kobuki base to move about the world, and the ROS framework to integrate the various sensors to build its world-view.

The design of the EMI robot is influenced by the design guidelines on the development of emotional agents, proposed by Shih et al. , with respect to the key ideas: Emotion, Design, Recognition, and Reaction [].

* Emotion: various design cues can be combined to formulate the emotional expression of the robot. EMI is designed with a larger spherical body (compared to its spherical head), a child-pitched voice and animated emotions in order to create its characteristic robot personality.
* Design and Recognition: The design and representation of emotion should generally be perceivable by humans. EMI adopts the projected head, an intuitive approach to realize emotional expressions. We have tested several designs including different materials for the projected head for the visual output. After several evaluations, we converged on a minimalist, cartoon-liked design, which is easily interpretable.
* Reaction: In order to elicit positive reactions from users, EMI was designed to be adorable, aligned with the concept of ‘kawaii’ (cute), a prominent aspect of Japanese popular culture.

# User Study

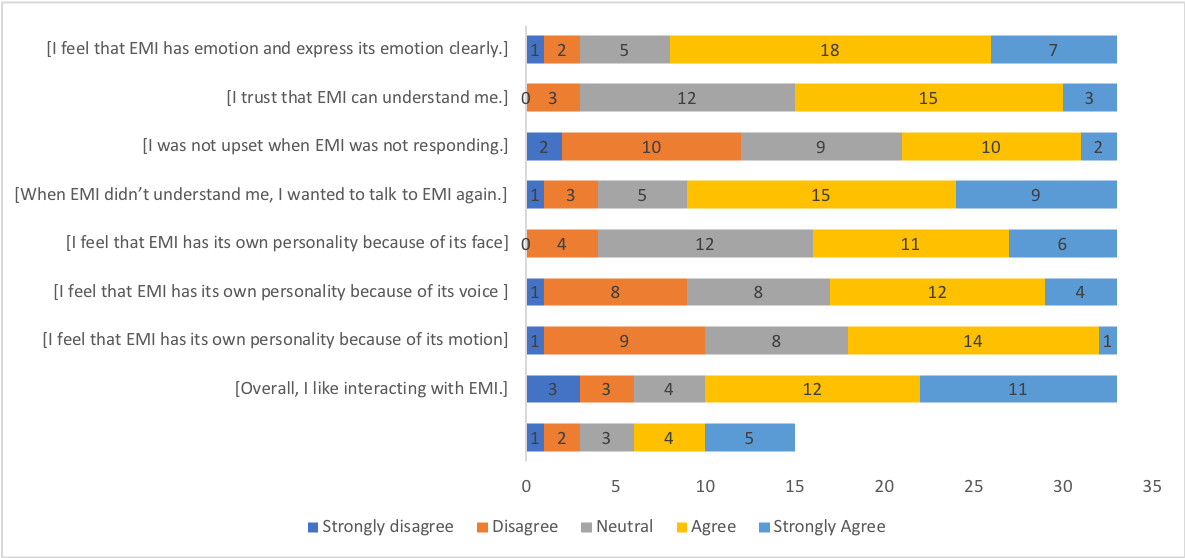
We have conducted an experiment to explore the effects of the emotional behaviors of EMI —projected face expression, voice, and motion—in EMI’s interaction with users. Our hypotheses were:

* The expressive behavior of EMI helps to increase tolerance and patience of users towards EMI.
* The emotional behavior of EMI helps to entice users to interact with EMI.

We prepared two mode of operations:

1. **Emotion Mode**. Users can ask EMI to demonstrate a specific expression, “EMI, can you smile”, “EMI, can you cry”, “EMI, can you sing a song”, “EMI, bye-bye”. Accordingly, EMI is prepared with the corresponding expressions related to smiling, crying, singing and kissing goodbye (shown in Figure []).

2. **Collaboration Mode** in which users can communicate with EMI, guiding EMI to navigate a maze. Users can talk to EMI through a microphone with a set of pre-defined commands, such as “EMI, turn left”, “EMI, turn right”, “EMI, go forward”, “EMI, go backward” and “EMI, stop”. EMI responds with verbal feedback that it understood the command. EMI also reacts to sensor inputs, such as, when it approaches obstacles (seen Table []).

We conducted an experiment at a demonstration booth at the Rakuten Technology Conference, which was open to the public. Our booth was set up within the “Kids Park” area, visited mainly by families with kids. An instruction manual listed the available commands and was prepared in both English and Japanese. All attendees could interact with the robot, but only those who agreed to the terms of our experiment and consented to the data collection policy are discussed in this paper. For those that agreed we recorded their interaction and had them fill out a survey after operating the robot. Users were invited to our booth’s control center, where they were given a microphone which was remotely linked to EMI for voice control. During the study, users were not required to follow any tasks, but were free to interact with EMI based on the available commands. When users looked confused about what to do, we would then give them some instructions.

In total, 33 users participated in the experiment, 20 Japanese speakers and 13 non-Japanese speakers. Their age ranges from 9 and 51 (M=31, SD = 15.2). Around half (51%) of participants have interacted with robots before, such as ‘Pepper’, a humanoid robot, designed for human interaction, typically as a store front receptionist.

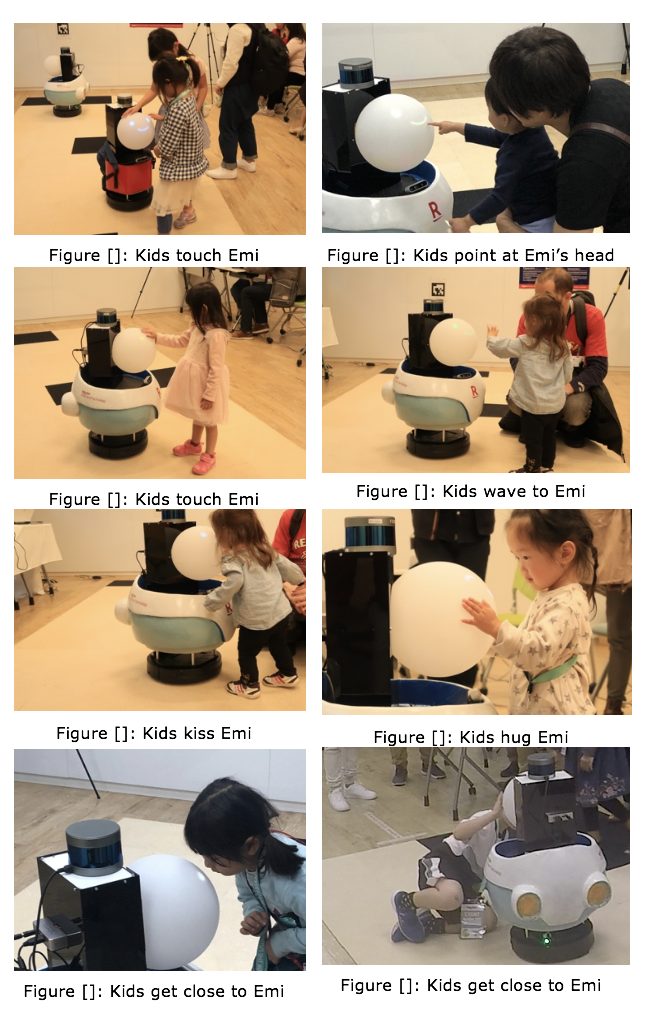
After their interaction with EMI, visitors were asked to complete a survey in order to provide feedback about various aspects of EMI. A 5-point Likert scale (“strongly agree, agree and neutral, disagree and strongly disagree”) was used to seek their agreement with various statements. A summary of the survey questions and results is illustrated in figure X. Questions requiring short answers, such as general comments about the robot or their impression of EMI’s personality, were also asked.

# Result and Discussion

## Here, we highlight our key findings from the survey and observations. We follow [] and converted our ordinal data into interval values and used statistical methods to find correlations. We used these equivalences: Strongly disagree = 1, Disagree = 2, Neutral = 3, Agree = 4, Strongly Agree = 5.

## EMI as an emotional robot

When asked if users agreed on “I feel that EMI has emotion and expresses its emotion clearly”, 76% of participants agreed or strongly agreed with this statement. One user specifically commented that the robot is “expressive and interactive”. Through a Mood’s Median Test we found these results was not gendered (Mood’s Median Test with p = 0.68).

Through our recordings we observed that users were often surprised or excited when they saw EMI’s positive reaction, making interjections like “aw” and “ooh”. When EMI was crying, some commented “かわいそう (poor thing)”. When EMI smiled users often smiled back at EMI and when EMI kissed good bye we observed a few of our younger users waving back at EMI and one young user kissed EMI back on its (projected) face. From both the survey data and the observation data, EMI’s expressions were is generally understood and accepted by users.

## Face is a relatively stronger factor compared to voice and motion.

We prepared three questions related to EMI’s personality as follows, “I feel that EMI has its own personality because of its face”, “… voice” and “… motion”, we obtained medians of 4,3,3, respectively. However, Mood’s median test indicated their differences were not statistically significant (p = 0.886).

Although the data does not provide strong evidence of the face’s significance, our observations indicate it was a strong factor. We observed many children being drawn to the round head, staring at it, pointing to it even touching, hugging and kissing it. Kids spent the most time responding the EMI’s face than the other factors (voice, motion).

## Tolerance and patience were mitigated

We hypothesized that when EMI makes mistakes user unease can be mitigated. In our experiment, there were two types of failure:

1. Failing to give a response when a user utters a command. This occurred, for example, when the wake-word was not recognized.

2. Apologizing for not being able to understand and asking to repeat the command. With phrases like, “sorry, I don’t understand” or “sorry, one more time”. This occurred when EMI reacted to the wake-word but the language unit failed to find a matching command.

The median response for “I was not upset when EMI was not responding” was 3. The responses were equally distributed, with 36% answered disagree or strongly disagree, while 36% agree or strongly agree, while the rest remained neutral. Subsequently, when asked “When EMI didn’t understand me, I wanted to talk to EMI again”, the median response was 4, with 73 % being agree or strongly agree. When EMI makes mistakes, users do indeed feel upset, but they were quite willing to give EMI a second chance.

Furthermore, from our observation, when EMI expressed its failure verbally, users appeared to be more tolerant, compared to when EMI does not give any response. When EMI gave no response, user’s attitude was more anxious and started becoming impatient. One user commented “I felt anxious when I waited for EMI’s response”. On the other hand, after EMI responded with “sorry I don’t understand”, most users showed disappointment, and most of the time, users repeated the commands. The tone when they repeated the commands appeared to be gentler. When EMI reacted successfully after several tries, users were relieved, and appeared happier. Finally, when asked, “I trusted EMI can understand me” the results showed that a majority agreed with this statement. We can conclude that users think that EMI is trying to understand them and so they are more patient and willing to try again to help EMI.

Our experiment was conducted in Japan and our visitors were composed of both native Japanese speakers and English speakers. Because of this, we additionally analyzed the possible difference between these groups. We found that English speakers (13 users) provided more positive responses than Japanese speakers (20 users). For the statement “I was not upset when EMI was not responding”, the median for English speaker vs Japanese speaker was 4 and 2 respectively. Interestingly, for the statement “When EMI didn’t understand me, I wanted to talk to EMI again” the median was 4 for both groups. The reason could be attributed to the fact that EMI’s speech recognition model was trained using American English phoneme from the **BLANK** database. So that it was better at understanding native-English speakers command compared to the non-native English speakers. Although the struggle from Japanese speakers to pronounce English makes them more frustrated with EMI, the tolerance and patience toward EMI was largely unaffected.

## Young kids treat EMI as a friend, while adults treat EMI as an infant

We observed that users of different age had different attitudes toward EMI. Kids had physical interaction with robot indicating friendship, such as hugging (shown in figure []). Apart from EMI’s face, they also touched EMI’s shell and camera, showing interest in EMI. Adults showed less physical interaction with EMI, but showed strong sympathy, saving EMI from moving too close to the edge of the course. When describing the personality, two users compared EMI to an infant. One couple stated that EMI is like their baby, when it moved to the edge, they got worried if it would get hurt, similar toe how they saw their young baby learning to move. “Cute”, “sweet” and “lovely” were common attributes given to EMI’s personality.

When asking “overall, I like interacting with EMI” 70% agree or strongly agree with this statement. Along with users commenting “I like it” and “I am looking forward for a future with EMI”, users were generally positive about EMI. The unique design decisions such as spherical head, projected animated face and voice created an effective social robot that promoted trust and tolerance.

*Aligning Results with Hypotheses*

To conclude, users perceived a relatively intimate social relationship with EMI by the emotional behavior design, which helps support our hypotheses. Hypothesis 1 was supported by the data – EMI draw attention from users. The social bond further encourages the interaction. Hypothesis 2 was supported by the data – Users appeared to be less upset with EMI failing the operation and be more generous about giving the retry.

# Conclusion

In this paper, we introduce EMI, an expressive, mobile, and interactive robot that is able to exhibit emotional behaviors, through the use of projected face expression, voice, and motion. We conducted an experiment to explore the effects of emotional and personal features in a human-robot interaction trial at a public event. We found that users were patient and tolerant of EMI even when it failed. This tolerance did not diminish even from users who were more prone to encountering these failures. From child users we observed the importance of the spherical head to draw engagement and EMI’s childish character helped adults empathize while commanding the robot. We think the design features of EMI provided a promising future for the design of social robots. We believe a commonly preferable animated emotion, a cute and characteristic voice together with expressive motion can help promote human-robot Interaction tasks.

Our experiment was conducted in a public setting, which enabled us to observe interesting and natural behaviors. However, as it is a public event, two limitations are that we were not able to designate a control group, and that our sample number is limited. In the future, we want to set a lab experiment, with two groups, one is performing a robot with the current configuration, one without the expressive features, to examine the same hypothesis. We will get more statistic support from that.