EMI: An Expressive Mobile Interactive Robot

Yuhui You
Rakuten Institute of Technology
yuhui.you@rakuten.com

Mitchell Fogelson
Rakuten Institute of Technology
mitchell.fogelson@rakuten.com

Kelvin Cheng
Rakuten Institute of Technology
kelvin.cheng@rakuten.com

Bjorn Stenger
Rakuten Institute of Technology
bjorn.stenger@rakuten.com

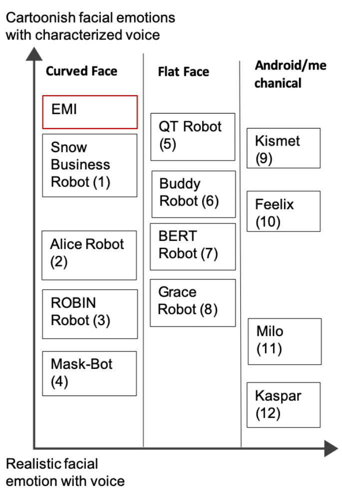


Figure 1: EMI and other robots mapped into different design features (1. [10] 2. [11], 3. [3], 4. [9], 5. [14], 6. [2], 7. [6], 8. [16], 9. [1], 10. [4], 11. [15], 12. [5])

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

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)

Abstract
In this paper, we explore how the emotional behavior of a robot affects interactions with humans. We introduce the EMI platform, an "Expressive, Mobile and Interactive robot", consisting of a circular diff-drive robot base equipped with a rear-projected expressive face, and omni-directional microphone for voice-interaction. We exhibited the EMI robot at a public event, in which attendees were given the option to interact with the robot and participate in a survey and observational study. The survey and observations focused on the effects of the robot's expressiveness in interactions with users of different ages and cultural backgrounds. From the survey responses, video observations and informal interviews we highlight key design decisions in EMI that resulted in positive user reactions.

Author Keywords
Human-Robot Interaction; Robot Design; Robot Tolerance; Emotional Robot

CSS Concepts
• **Computer systems organization~Robotics**

Introduction
In the study of Human-Robot Interaction (HRI), there is an interest in developing solutions for social interactions between humans and autonomous robots. As robots operate in human inhabited spaces, the

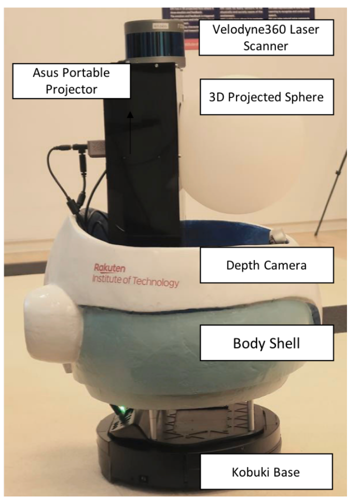


Figure 2: EMI's architecture

environment can become difficult to interpret and complex to navigate. One approach in managing 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. These interactions require the robot to enlist help from an inattentive human. In human-human interaction a combination of verbal and physical actions is used to draw a potential helper's attention, who typically will respond verbally to the request. For someone to agree to assist, several features seem to be linked: tone of voice, size of request, politeness, status and more [17]. For human-robot interaction, we propose EMI – an expressive, mobile and interactive robot – to tackle this highly coupled challenge in an effort for it to work in our office environment. EMI was designed to be an emotional social robot that combines facial animation, voice, and motive expressiveness, in order to better enlist human assistance when necessary. We had an opportunity to test our robot design at a public exhibition that was held at our organization. During the event our aim was to better understand the aesthetic design decision we made with respect to the following questions: Does EMI's expressiveness increase tolerance and patience of users communicating with the robot? Does EMI's expressiveness make it more engaging? We also wanted to see how the responses differed between children and adults when interacting with the robot. We believe our observations will help contribute to improving future social robot design.

EMI Robot

We developed a new robot called EMI in an effort to better address the highly coupled HRI challenges for a robot operating in an office space. We examined various robot designs in the market and academia which influenced our design. They are categorized into different design space, which is shown in Figure 1.

EMI's physical structure is built upon the Turtle Bot 2E robot and is equipped with a 3D laser scanner and Intel RealSense D435 camera. An Asus projector and a spherical lamp shade are used to create EMI's face. Finally, a standard omni-directional USB microphone is used to receive speech input. EMI uses an onboard Intel NUC computer and the ROS framework [13] to integrate the subsystems and the various sensors to build its world-view.

We crafted the facial animation, voice and motive expressiveness to form an identifiable personality. EMI is able to express different emotions using the projector, in conjunction with voice output, shown in Figure 3. Similar to smart speaker devices, EMI's speech recognition is equipped with wake-up word detection and command recognition using the ASpIRE speech database [7] and Kaldi framework [12] for speech transcription. It uses natural language understanding to convert the speech transcription to intent categories. EMI is built using the differential drive Kobuki base and can be controlled by temporal velocity commands to the motors to create movement.

The aesthetic design of the EMI robot is influenced by the design guidelines on the development of emotional agents, proposed by Ruud et al., with respect to the key ideas: Emotion, Design, Recognition, and Reaction [8].

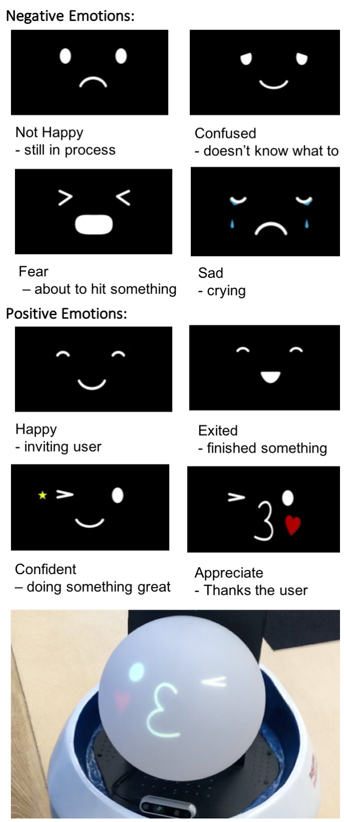


Figure 3: EMI's expression design and actual visual appearance.

- 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's spherical head and minimalist, cartoon-liked design creates an easily interpretable face.
- Reaction*: The design should elicit positive reactions.* EMI was designed to be adorable, aligned with the concept of 'kawaii' (cute), a prominent aspect of Japanese popular culture.

User Study

We conducted an experiment to explore the effects of the emotional responses of EMI while interacting with users. Our hypotheses were:

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

We categorized the operation into two types. Since the input is user's speech, user can ask Emi to response based on the speech commands.:

1. **Emotion Operation**: Users can use speech commands to ask EMI to demonstrate a specific expression, saying commands such as "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.
2. **Collaboration Operation**: We have designed a maze operation scenario for user to collaborate with Emi. Users can give directional commands moving EMI, 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 and the corresponding motive output. EMI also reacts to sensor inputs, such as, when it approaches obstacles (seen Table 1).

The experiment took place at a demonstration booth at the Rakuten Technology Conference, which was open to the public seen in figure 4. Our booth was set up within the "Kids Park" area, visited mainly by families with children. 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 who agreed, we recorded their interaction and had them fill out a survey after operating the robot. During the study, we place Emi on our designed maze area. users were not required to follow any tasks and were free to interact with EMI based on the available commands, so they can either help Emi walk through the maze, or they can just simply ask Emi to show some emotion. They were not required to finish the maze task. If they didn't want to continue, they can drop off at any time. When users looked confused about what to do, we would give them some instructions. After their interaction with EMI, visitors completed a survey about various aspects of the experience. 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 5. Questions requiring short answers, such as general comments about the robot or their impression of EMI’s personality, were also asked.

In total, 33 users participated in the survey, 20 Japanese speakers and 13 non-Japanese speakers. Their age ranges from 9 and 51 (M=31, SD = 15.2). Around half (51%) of the participants have previously interacted with robots, such as ‘Pepper’, a humanoid robot, designed for human interaction, typically as a store front receptionist. Besides that, we also had observation from the whole experiment period, which came from those who participated in the survey and those who didn’t, including participators’ children, who are not capable of answering questions.

Result and Discussion

Here we highlight our key findings from the survey and observations. We 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 with the statement, “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 are 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 and observation data, EMI’s expressions were 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 ($p = 0.89$).

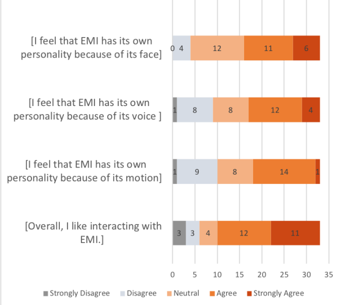
Although the data does not provide strong evidence of the face’s significance, our observations indicate it was a strong factor. We observed many children being drawn to the round head, staring, pointing, touching, hugging and kissing it. Kids spent more time responding to EMI’s face than other features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

Tolerance and patience were mitigated

We hypothesized that EMI mitigates user discomfort when errors occur. In our experiment, there were two types of failure: **1. Failing to give a response when**

Figure 4: The demonstration booth. EMI saw the black block as obstacles.

Environment	Voice	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		



a user utters a command. This occurred 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 recognized the wake-word but the NLU failed to find a correlating intent from the 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 and 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.

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. Although the struggle from Japanese speakers to pronounce English makes them more frustrated with EMI, the tolerance and patience toward EMI was largely unaffected supporting hypothesis 1.

Young kids treat EMI as a friend, while adults treat EMI as a child

We observed that users of different age groups had unique attitudes toward EMI. Kids tended to enjoy physically interacting with the robot in friendly ways, such as hugging EMI's face (shown in figure 6). Apart from EMI's face, they also touched EMI's shell and camera, showing interest in the embodiment. EMI is approximately the same size as a child age 4-5 ~0.9m and saw eye to eye with some of the child users.
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 experience with EMI, two adult users compared EMI to a child. One couple stated that EMI is like their baby, when it moved to the edge, they got worried if it would get hurt, similar to how they saw their young baby learning to move. "Cute", "sweet" and "lovely" were common attributes given to EMI's personality.
When asked about the statement, "Overall, I like interacting with EMI", 70% agree or strongly agree. Along with users commenting they enjoyed the experience and "looking forward for a future with EMI." The 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

Users perceived a relatively intimate social relationship with EMI supporting our hypotheses.
Hypothesis 1 -- The emotional behavior of EMI helps to increase tolerance and patience of users towards EMI -- was supported by the data showing that a tolerance did not decrease between the Japanese speakers vs English

speakers even though the Japanese speakers felt greater frustration with the interaction.
Hypothesis 2 – The emotional behavior of EMI helps to entice users to interact with EMI-- was supported by the observations of adult and child responses while interacting with the robot.

Conclusion

In this paper, we introduced EMI, an expressive, mobile, and interactive robot that is able to exhibit emotional behaviors, through the use of projected face, 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 with EMI even when it failed. This tolerance did not diminish even from users who were more prone to encountering failures. From children we observed the importance of the spherical head to draw engagement and EMI's childlike character helped adults empathize while commanding the robot. We think the design features of EMI provided interesting insights for future social robots. We believe a commonly preferable animated emotion, a cute and characteristic voice together with expressive motion can help promote human-robot interactions.


Figure 6: Children interacting with EMI.

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 was limited. As future work, we intent to conduct a control study with two conditions, one with the robot in its current design configuration, and a control condition without the emotional and expressive features, in order to further validate our hypotheses empirically.

Acknowledgement

This research was supported by different teams in Rakuten Institution of Technology, robotics team including Mathilde Chachura and Wilson Colin working on robot hardware and software development, speech team including Federico Ang, Congying Zhang, Mausam Jain and Yuan Liang. We also want to thank for Aki Nakazawa for assistance with the design.

References

[1] Cynthia Breazeal. 2003. Emotion and sociable humanoid robots. *International journal of human-computer studies* 59.1-2: 119-155. https://www.sciencedirect.com/science/article/pii/S1071581903000181

[2] BLUE FROG ROBOTICS. BUDDY - The Emotional Robot. Retrieved December 22, 2019 from https://buddytherobot.com/en/buddy-the-emotional-robot/.

[3] Karsten Berns and Zuhair Zafar. 2018. Emotion based human-robot interaction. *MATEC Web of Conferences*. Vol. 161. EDP Sciences, 2018. https://www.matec-conferences.org/articles/matecconf/abs/2018/20/matecconf_erzr2018_01001/matecconf_erzr2018_01001.html

[4] Lola D Cañamero. 2002. Playing the emotion game with feelix. *Socially intelligent agents*. Springer, Boston, MA. 69-76. https://link.springer.com/chapter/10.1007/0-306-47373-9_8

[5] Kerstin Dautenhahn, Chrystopher L. Nehaniv, Michael L. Walters, Ben Robins, Hatice Kose-Bagci, N. Assif Mirza, and Mike Blow. 2009. KASPAR–a minimally expressive humanoid robot for human–robot interaction research. *Applied Bionics and Biomechanics* 6.3-4: 369-397. https://www.hindawi.com/journals/abb/2009/708594/abs/

[6] Adriana Hamacher, Nadia Bianchi-Berthouze, Anthony G. Pipe, and Kerstin Eder. 2016. Believing in BERT: Using expressive communication to enhance trust and counteract operational error in physical Human-robot interaction. *25th IEEE international symposium on robot and human interactive communication (RO-MAN)*: 493-500. https://ieeexplore.ieee.org/abstract/document/7745163/

[7] Mary Harper. 2015. The automatic speech recognition in reverberant environments (ASpIRE) challenge. *2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*. IEEE. https://ieeexplore.ieee.org/abstract/document/7404843

[8] Ruud Hortensius, Felix Hekele, and Emily S. Cross. 2018. The perception of emotion in artificial agents. *IEEE Transactions on Cognitive and Developmental Systems* 10.4852-864. https://ieeexplore.ieee.org/abstract/document/8341761

[9] Takaaki Kuratate, et al. 2011. "Mask-bot": A life-size robot head using talking head animation for human-robot communication. *2011 11th IEEE-RAS International Conference on Humanoid Robots*. IEEE. https://ieeexplore.ieee.org/abstract/document/6100842

[10] Mike's Robot Lab. Snow Business Solution Robot. Retrieved December 22, 2019 from https://mikes-robot-lab.myshopify.com/collections/frontpage/products/snow-business-solution-robot.

[11] Mike's Robot Lab. Alice Business Solution Robot with 23.8 Inch Screen. Retrieved December 22, 2019 from https://mikes-robot-lab.myshopify.com/collections/frontpage/products/alice-business-solution-robot-with-23-8-inch-screen

[12] Daniel Povey, Arnab Ghoshal, Gilles Boulianne, Lukas Burget, Ondrej Glembek, Nagendra Goel, Mirko Hannemann et al. 2011. The Kaldi speech recognition toolkit. *IEEE 2011 workshop on automatic speech recognition and understanding*. No. CONF. IEEE Signal Processing Society. https://infoscience.epfl.ch/record/192584

[13] Quigley, Morgan, Ken Conley, Brian Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y. Ng. 2009. ROS: an open-source Robot Operating System. *ICRA workshop on open source software*. Vol. 3. No. 3.2. https://www.willowgarage.com/sites/default/files/icraoss09-ROS.pdf

[14] Francisco J. Rodriguez-Lera, Leandro Gomes, Pouyan Ziafati, Aida Nazarikhorram, Andrea Stefanetti, and Anne-Marie Schuller. 2018. Emotional robots for coaching: Motivating physical rehabilitation using emotional robots. *Proc. Personal Robots for Exercising and Coaching (PREC)*: 1-7.

[15] Robots4Autism. Meet Milo!. Retrieved December 22, 2019 from https://robots4autism.com/milo/.

[16] Reid Simmons, Dani Goldberg, Adam Goode, Michael Montemerlo, Nicholas Roy, Alan C. Schultz, Myriam Abramson, Ian Horswill, David Kortenkamp, and Bruce Maxwell. 2003. *Grace: An autonomous robot for the AAAI robot challenge*. CARNEGIE-MELLON UNIV, PITTSBURGH, PA.

[17] Vasant Srinivasan , and Leila Takayama. 2016. Help Me Please: Robot Politeness Strategies for Soliciting Help from People. *Conference on Human Factors in Computing Systems - Proceedings*, pp. 4945–55. https://dl.acm.org/citation.cfm?id=2858217