J Multimodal User Interfaces
DOI 10.1007/s12193-009-0028-2

ORIGINAL PAPER

When my robot smiles at me

Enabling human-robot rapport via real-time head gesture mimicry

Laurel D. Riek · Philip C. Paul · Peter Robinson

Received: 15 April 2009 / Accepted: 9 November 2009

Abstract People use imitation to encourage each other during conversation. We have conducted an experiment to investigate how imitation by a robot affect people's perceptions of their conversation with it. The robot operated in one of three ways: full head gesture mimicking, partial head gesture mimicking (nodding), and non-mimicking (blinking). Participants rated how satisfied they were with the interaction. We hypothesized that participants in the full head gesture condition will rate their interaction the most positively, followed by the partial and non-mimicking conditions. We also performed gesture analysis to see if any differences existed between groups, and did find that men made significantly more gestures than women while interacting with the robot. Finally, we interviewed participants to try to ascertain additional insight into their feelings of rapport with the robot, which revealed a number of valuable insights.

Keywords Affective computing · Empathy · Facial expressions · Human-robot interaction · Social robotics

1 Introduction

The expression of empathy is a key aspect of human-human social communication that allows people to experience and understand what others are emotionally conveying [19]. One of the most basic forms of expressive empathy is known as emotional contagion [8, 9], where an observer mimics the behavior of a target, and by virtue of that mimicry, comes to experience an emotional state similar to that of the target [3].

Facial expression and head gesture mirroring are common forms of empathic conveyance that typically include head nodding, laughing, eyebrow raising, smiling, and so on. This mirroring is so vital to emotional communication that if an individual's ability to mirror others is physically blocked, that individual will actually be impaired in their ability to identify emotions [14].

Given the importance of facial mimicry in human-human communication, we wondered if it might also be important in human-machine communication. In particular, might a conversational robot that mimics a few low-fidelity expressions and head gestures in real-time create a more satisfying interactive experience for people? To address this question, we built a head gesture mimicking robot named Virgil and performed a user study to explore this question.

2 Head-gesture mimicking robot

2.1 Motivation

From a technological perspective, a number of virtual and robotic systems have been built that incorporate real-time facial expression mimicking. For virtual avatars representing people, Kang et al. [11] provides a thorough survey. Gratch et al. [6] developed the Rapport Agent, an intelligent virtual agent that tries to engender rapport with users who interact with it. Baileson and Yee also built a intelligent virtual

---

This research is sponsored by the Qualcomm Research Studentship.

L.D. Riek (✉) · P. Robinson
Computer Laboratory, University of Cambridge,
Cambridge CB3 0FD, UK
e-mail: Laurel.Riek@cl.cam.ac.uk

P.C. Paul
Department of Engineering, University of Cambridge,
Cambridge, UK

Fig. 1 Virgil, our conversation robot

agent that subtly mimicked people’s head gestures in real-time, and found this made the agent more persuasive [2].

In robotics, some researchers have explored real-time conveyance of facial expressions, gaze, and head gestures on physical avatars, which are tele-operated robots intended to represent remotely located users; see Riek [16] for a survey of this work. Also, work has been done on autonomous robotic platforms that have humanlike or animal-like appearances that convey expressions in real-time, see Walters [22] for a thorough survey. Nadel et al. compared how participants respond to affective facial displays on a robot compared with a human [13].

From a psychological perspective, it is well understood that humans and some non-human mammals can convey empathetic responses via involuntary facial mimicry. This is mimicry that does not involve a cognitive dimension and is quickly processed, usually within one second or less [19]. In that vein, we created a naive head gesture mimicking robot capable of mirroring a human’s head gestures in real-time.

2.2 Platform description

We chose to use the WowWee Alive Chimpanzee Robot [25], which we have named Virgil (see Fig. 1). Practically, this robot was selected because it was inexpensive and easily modifiable. However, it was also selected because apes can be extremely empathetic creatures [19], and are thus a natural platform to try for facial mimicry.

Virgil has a total of 18 degrees-of-freedom (DOFs). Its eyes have 4 DOFs (up/down/left/right), eyebrows 2 DOFs (up/down), its lower jaw 2 DOFs (up/down), its upper lip 2 DOFs (up/down), and its head 8 DOFs (roll/pitch/yaw). Out of the box the robot can operate fully autonomously or be tele-operated via remote control.

While the robot can be operated via remote control, such control is nowhere near sufficient to facilitate real-time movement commands. Thus, we modified the robot so we could directly control all its motors via an Arduino microcontroller [1].

2.3 Software

Figure 2 contains a graphical representation of the end-to-end system. In short, the software components were as follows:

Face tracking To capture a user’s head movements, we used a face tracking visualization program developed by Harrison [7], which uses the OpenCV Haar Cascade classifier to identify faces [24].

Movement selection After receiving a set of location and pose coordinates of the head from the face tracker, the software decides what the appropriate movement response should be from the robot. This decision depended on which experimental condition the participant was assigned to. (See Sect. 5 for a more detailed description of each conditions.) This software was written using the Processing programming language [15].

Movement commands After deciding what move to make, the command is sent to the appropriate output on the Arudino via the Firmata protocol [21].

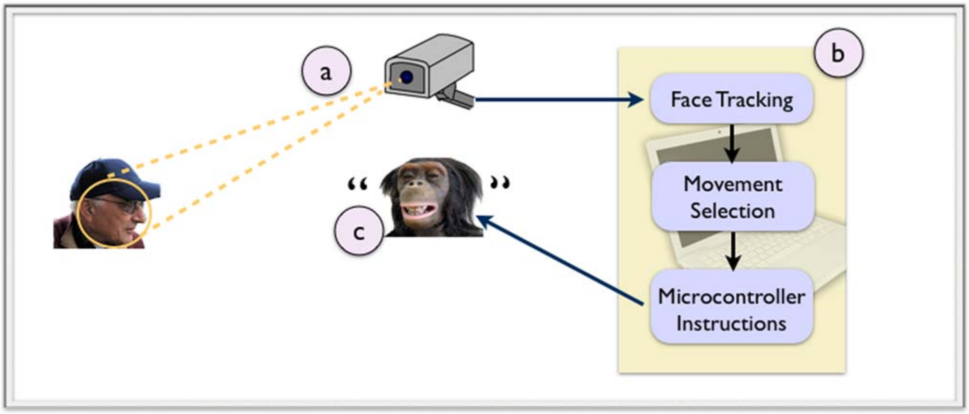


Fig. 2 The end-to-end system. (**a**) The video camera captures the user’s head movements. (**b**) The OpenCV face tracker extracts the facial movements, the movement selector decides which way to move the robot, then microcontroller instructions are transmitted to the robot. (**c**) The robot moves

3 Pilot study

Before our end-to-end system was complete, we wanted to explore the ideas of how non-verbal communicative gestures of empathy might alter people's affective states, as well as get some initial feedback on our system design. So we ran a pilot study [18] with Virgil and a user interacting one-on-one in our usability lab. We asked participants to tell a non-personal and personal story to the robot.

The pilot study had an experimental and control condition. In the experimental condition we wizard-of-oz controlled the robot to make empathetic head gestures (nodding) and mouth movements (open mouth in surprise). In the control condition, the robot acted autonomously in a random manner.

Following the study, we asked participants to complete a written questionnaire that asked them to rate their interaction with the robot. These questions were 5-point Likert scale items such as "I think Virgil could be a friend of mine," and "Virgil recognized my feelings and emotions appropriately for the situation". This resulted in each participant having an overall interaction satisfaction score. As we predicted, participants in the experimental condition rated the interaction more favorably than those in the control condition [18]. However, because our two experimental conditions were so extremely different, this result is not surprising. Therefore, we decided we needed a more fine-grained approach to our experimental conditions, which we will discuss in Sect. 5.

4 Research questions

We wanted to understand the degree to which participants might like a robot to mimic their head gestures in real-time. Due to the noisiness of the robot's motors and lack of movement precision, we wondered if a fully mimicking robot would be perceived as annoying. We also wanted to know if a robot that mimics just some head gestures some of the time people might encourage people to nod in response to it. In addition, we wondered what the effects might be of a robot that just blinks at people. Finally, we wondered if people made body gestures more frequently in some conditions as opposed to others.

Thus, we designed an experiment to test three head-gesture mimicking conditions:

Full: The robot mimics all head gestures.
Partial: The robot mimics up/down nodding gestures only.
None: The robot blinks periodically.

For all the robot's head movements, responses were in real-time to the participants' movements without any noticeable delay. Also, for any left/right movements in the full mimic condition, the robot's head movements were mirroring the participants. So if the participant moved their head to the left, the robot moved its head to the participant's left (the robot's right).

Thus, our hypotheses are as follows:

- Participants in the full head gesture mimicking condition will rate their interaction as more positive than participants in the partial head mimicking and non-mimicking conditions.
- Participants in the partial head gesture mimicking condition will rate their interaction as more positive than participants in the non-mimicking conditions.

These hypotheses are motivated by the idea that people will appreciate a robot that mimics their head movements in real-time the most, followed by a robot that nods, followed by the blinking robot. This motivation comes from the work of Sidner et al. [20], who ran an experiment with a penguin robot capable of making head nods and detecting human head nods. They found a similar tertiary ranking—people nod more often when conversing with a robot that nods deliberately in response to their nods, compared with one that merely recognizes their nods, compared with one that doesn't nod at all. Similarly, Kanda et al. [10] found that people rated interaction with a robot more highly when it made appropriate temporal-cooperative behaviors (such as nodding at the right time in the conversation) and spatial-cooperative behaviors (such as gaze and gesture sharing) compared with a robot that did not employ such behaviors.

5 Experimental design

The experimental design is a 3 × 1 between-subjects factorial design involving three levels of head gesture mimicry: full head gesture mimicking (up/down/left/right movement), partial head mimicking (up/down nodding), and non-mimicking (blinking).

For this experiment, the robot was fully autonomous across all three of the conditions.

5.1 Materials

5.1.1 Instruction sheet and consent forms

We prepared an instruction sheet for participants that contained and overview of the study, a description of the procedure, and standard consent protocols.

Participants also received a consent form to be videotaped. Participants could choose to not be videotaped and still participate in the experiment.

5.1.2 Robot information sheet

In order to mitigate some of the emotions we expected people to experience upon first encountering the robot (surprise, fear, discomfort, etc.), we prepared a robot description sheet that had a picture of the robot with a brief textual description of its capabilities. This included information about how the robot would move (only its head and eyes), and that it would be silent aside from its motor movements.

5.1.3 Tasks

The experiment consisted of two verbal tasks. The first task asked participants to describe the route they took to the laboratory that day. We gave participants this task first because navigation is something that requires non-emotional, yet cognitive, thinking. Our intention was to put participants at ease with the robot and with the experimental settings.

The second task was: "Next, please tell Virgil about your first memories of Cambridge—people you met, things you saw, foods you ate, etc. Please be as descriptive as possible." We selected this task because we wanted something that was common across all participants regardless of their job, age, or cultural background. We also wanted a task that was both emotionally salient and personal so participants would feel invested in discussing it with a robot.

5.1.4 Post-evaluation questionnaire

In order to measure how satisfied participants were interacting with the robot, we used a modified version of the Interactant Satisfaction Survey developed by Kang et al. [11]. This survey is a fifteen item, 8-point Likert scale that measures the social attraction toward and emotional credibility of conversation partners. (See Table 1.) For human-human communication this measure has high internal consistency (Cronbach's $\alpha = 0.90$) [12].

5.1.5 Post-evaluation interview

At the end of the experiment, participants took part in a semi-structured interview with the experimenter. They were asked about their first impressions of Virgil, how they felt talking to a robot, and if they felt like the robot made an amicable conversation partner. The experimenter asked follow-up questions when appropriate, and encouraged participants to elaborate as much as possible.

5.2 Participants

We recruited 12 participants to participate in our study. Participants were recruited via email, word of mouth, and Facebook. They were told they would be participating in a study where they would be asked to talk to a robot. Participants were remunerated for their participation.

Six participants were female and six were male. Their ages ranged from 19–70 years old ($M = 32.1$, s.d. $= 14.5$). Six participants were born in England, two in China, and one each from India, Spain, Pakistan, and the United States. Eight participants were students of various types (PhD, undergraduate, and sixth form), one was retired, one was a stay-at-home parent, and two were managers. In terms of technological experience, two participants considered themselves novices, seven considered themselves average, and three considered themselves experts.

Table 1 Modified version of the Interactant Satisfaction Survey [11]

Interactant Satisfaction
Social Attraction Items
I think Virgil could be a friend of mine.
I would like to have a friendly chat with Virgil.
We could never establish a personal friendship with each other.
Virgil just wouldn't fit into my circle of friends.
Virgil would be pleasant to be with.
I don't care if I ever get to interact with Virgil again.
Virgil recognizes my feelings and emotions.
Virgil expresses feelings and emotions appropriately for the situation.
Emotional Credibility Items
Virgil uses feelings and emotions to create or organize thinking.
Virgil uses feelings and emotions to make a decision or judgment.
Virgil uses feelings and emotions to facilitate problem solving and creativity.
Virgil responds appropriately to positive and negative emotions.
Virgil understands complex feelings.
Virgil knows how to control its feelings and emotions effectively.
Virgil handles others' feelings and emotions sensitively and effectively.

5.3 Procedure

Participants were randomly assigned to one of the three experimental conditions: full head gesture mimicking, partial head gesture mimicking, and non-mimicking. Four participants were in each group. After reading the study instructions and signing a consent form, participants were brought to an office and told to be seated in a chair in front of Virgil (see Fig. 3). The position of the chair, robot, and camera were fixed, and remained constant across all participants.

After a participant was seated, the experimenter left the room, and verbally gave them the first task from behind the door. The experimenter then closed the door and left the participant alone with the robot to complete the task. The participant verbally signaled to the experimenter when they were finished, and the same procedure was repeated for the second task. Following the tasks, the experimenter entered the room and turned the robot off. The participant was re-seated at a desk away from the robot and was handed the written questionnaire to complete—the experimenter again left the

Fig. 3 The experimental setup. A participant sits in a chair facing the robot, about one meter away from it. The camera and chair position remain fixed across all participants

room. After the participant completed the questionnaire, the experimenter re-entered the room and performed a qualitative interview with the participant.

6 Measures

6.1 Interactant Satisfaction Survey

The Interactant Satisfaction Survey is a summative measure that provides an overall satisfaction score that ranges from 0–120. To calculate the score, three questions must be recoded (#3, #4, and #6) to be positively biased. Then the responses can be summed. Because this measure provides a total numeric score, we used parametric statistics to analyze our results.

6.2 Gesture analysis

In addition to the Interactant Satisfaction Survey, our other primary measure was gestural analysis of the video data. We coded the gestures as follows:

Head gestures For coding the head gestures people made during their interaction with the robot, we looked at the types of head nods people made as well as the number of times they smiled at the robot. For head nods, we measured discrete head movements in the left/right and up/down direction. These might have been a nod, tilt, or roll. Nods, tilts, and rolls along the left/right direction were all counted as "Left/Right Nods", and any of these gestures in the up/down or forward/back direction were counted as "Up/Down Nods".[1] (See Fig. 5.)

Body gestures For body gestures, we analyzed foot and leg movements, general torso movements, shrugs, and leaning movements. Foot and leg movements included any discrete motions we could observe. These included kicking and rotating the foot, as well as bouncing either the foot or leg. For leaning movements, some participants spent a lot of time leaning forward to talk to the robot.

Hand gestures We also examined hand gestures. These included discrete gestures used while speaking, fidgets, and hand wringing. For fidgets, we counted discrete movements such as pulling the sleeve, grabbing the arm, rubbing the leg, moving the hand up and down the leg. We considered hand wringing to be a special type of fidget, and for this we counted discrete hand movements, such as hand squeezes, fingers opening and closing, and hand or finger pulls.

Sound Finally, we were interested in the number of times people laughed while speaking to the robot, and if they said 'Hi' the first time they met the robot.[2]

7 Results

7.1 Manipulation checks

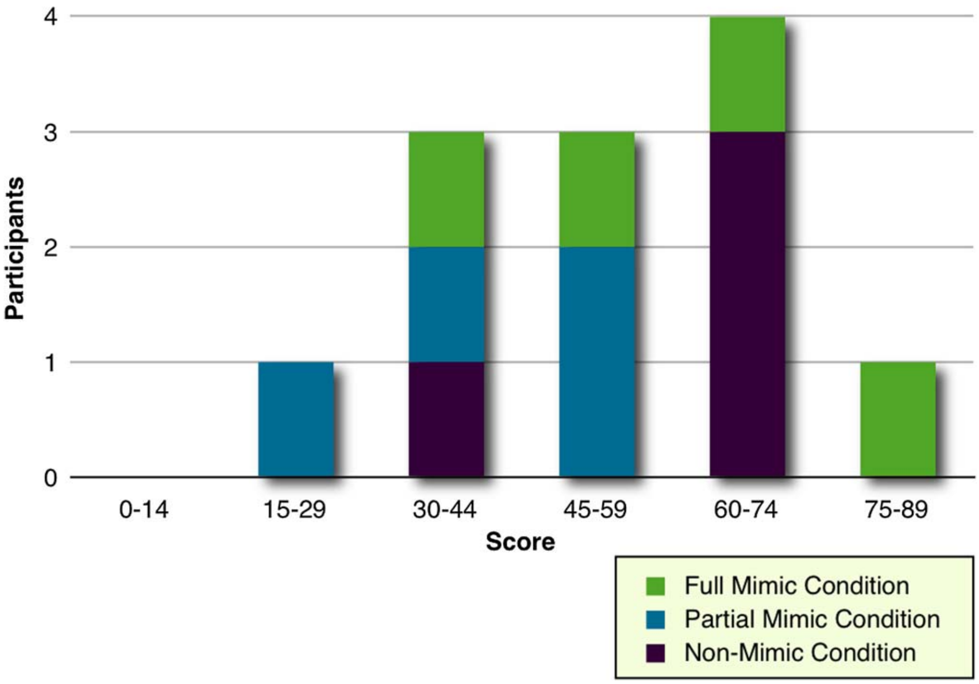
Two participants had seen the robot a few times before the experiment, but never when it was turned on. The other ten participants had never seen the robot before.

---

[1] We included forward/back head gestures under the "Up/Down Nod" category because given the angle our camera was set at it was difficult to distinguish the two.

[2] We were also interested in speech speed, as two participants spoke extremely slowly to the robot, as though they were speaking to a speech recognition system, and interestingly, they both later reported talking to a robot felt like talking to an automated telephone system. However, we did not perform any formal speech analysis for this study.

Fig. 4 The distribution of scores on the Interactant Satisfaction Survey for each of the three conditions. Scores ranged from 15–80 (S.d. = 18.6)



Two participants knew the experimenter was researching affective computing (but not robotics), and two participants knew the experimenter was researching robotics (but not affective computing). No participants knew what the study was testing, or that there were multiple conditions being tested.

7.2 Interactant satisfaction results

The scores on the Interactant Satisfaction Survey ranged from 15–80 (s.d. = 18.6). See Fig. 4 for a visualization of the distribution of the raw scores.

For Interactant Satisfaction, ANOVA results did not show a statistically significant effect between the three conditions ($ps > .05$). This result may be due to the fact that we had too small a number of participants per group. Another possibility is that this measure is not well-suited toward one-way communication with a robot. In any case, further experimentation is needed.

Thus, we fail to find support for our first two hypotheses —we found no statistically significant difference between the three conditions in terms of Interactant Satisfaction.

7.3 Gesture analysis results

7.3.1 Quantitative results

Figure 5 contains a table detailing the mean number of gestures participants made per category (Head, Body, Hand, or Sound) and per condition (Full, Partial, or Non-Mimicking).

Condition A one-way ANOVA on total gestures made between the three conditions showed no statistically significant effect.

Age Dividing the participants up into equally distributed age groups (<19, 20–29, 30–39, 40+), we performed a one-way ANOVA to see if age played a role in total gestures made while interacting with the robot, but the results were not significant.

Expertise A one-way ANOVA comparing expertise (Novice, Medium, and Expert) and total gestures made while interacting with the robot was not significant.

Country of birth Given our sample was roughly half British, we decided it would be reasonable to split people into one of two country groups, British and non-British, to see if there were any differences between groups. With this deliniation, we did not find a statistically significant difference in total gestures made.

Gender In the experiment, male participants made significantly more gestures ($M = 104.83$, $SE = 15.0$) compared to female participants ($M = 53.40$, $SE = 12.54$). This difference was significant, $t(9) = -2.56$, $p < .05$, and it represents a large effect size, $r = .65$.

7.3.2 Qualitative results

Analysis of participants' gestures and speech in the video recordings revealed several interesting features of their interactions with the robot.

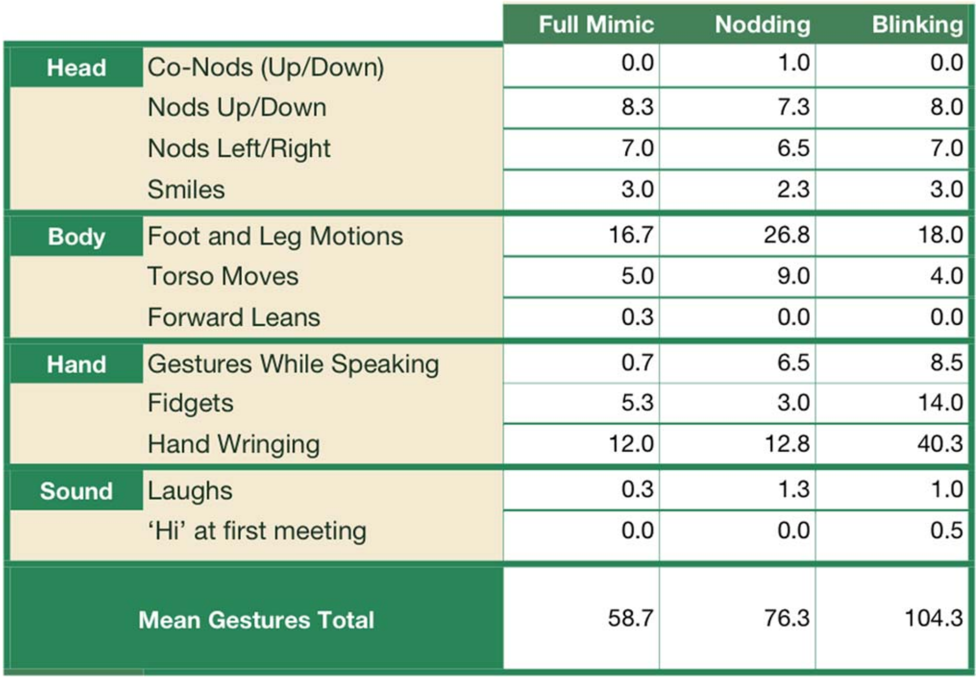


Fig. 5 The mean number of gestures participants made per condition during their interaction with Virgil. The conditions were full-mimicking, partial-mimicking (nodding), and non-mimicking (blinking)

		Full Mimic	Nodding	Blinking
Head	Co-Nods (Up/Down)	0.0	1.0	0.0
	Nods Up/Down	8.3	7.3	8.0
	Nods Left/Right	7.0	6.5	7.0
	Smiles	3.0	2.3	3.0
Body	Foot and Leg Motions	16.7	26.8	18.0
	Torso Moves	5.0	9.0	4.0
	Forward Leans	0.3	0.0	0.0
Hand	Gestures While Speaking	0.7	6.5	8.5
	Fidgets	5.3	3.0	14.0
	Hand Wringing	12.0	12.8	40.3
Sound	Laughs	0.3	1.3	1.0
	'Hi' at first meeting	0.0	0.0	0.5
Mean Gestures Total		58.7	76.3	104.3

Co-nodding Two participants, both in the partial mimicking (nodding) condition, co-nodded with the robot. In other words, the participant nodded, the robot nodded in response, and then the participant nodded to acknowledge the robot's nod. This ability for the robot to influence the user's behavior was something entirely unexpected, and we hope to do follow-up studies to explore this result.

Posture A few participants adopted unexpected postures when interacting with the robot. Some participants leaned forward to talk to the robot (see Fig. 3 for an example of this leaning). Others adopted a slightly slouched posture. While none of these postures were so extreme as to affect our face tracker, it is worth noting that even when seated people may move around in a way that may render a face tracker completely useless. A wide-angle camera lens may help address this potential issue.

Deliberate behaviors While we were not aiming to investigate the tendency of participants to "play the system" when interacting with affect-sensitive machines [17], we did note a few interesting things with regard to deliberate behaviors. One participant in the non-mimicking condition spoke extremely slowly to the robot, and later reported that talking to the robot was like "talking to an automated machine". (e.g., an automated telephone service). A few other participants also spoke in a strange manner to the robot on occasion during their interaction. We have no way of knowing why people altered their speech in this manner, but we plan to explore this in future research. In particular we are interested to know whether it was due to nervousness, individual differences in expression, or attitudes regarding hatred of speech-recognition systems.

Problematic gestures In this experiment, we chose to base our robot's movements solely on the basis of head gestures made by participants. However, this turned out to be problematic, because some participants never moved their head despite making significant movements with other parts of their body. For example, one participant in the partial head gesture mimicking condition never nodded his head, but moved his foot a total of 70 times. Thus, it may be important for us to look at more than just head movements in future work.

7.4 Interview results

As with our pilot study [18], we received a wide range of responses from participants during our post-experimental interview. We plan to incorporate some these responses into the design of both future interactive robots and future experiments.

7.4.1 Interaction was machine-like

A few participants commented that talking to the robot was like talking to a machine: "Once I got used to what [the robot] looked like visually, I almost switched [it] off… the fact that I was talking to a robot [not] an animated object in front of me—I almost treated it like it was an answering machine at the end of a phone…I don't think I got anything more from the experience because it was a robot. You could probably just put me in front of the camera and I would have had a similar kind of reaction" [Non-mimic condition].

Another participant made a similar remark about how talking to the robot was like talking to an automated machine, and then remarked "You could talk to a statue and get as much of a response" [Partial mimic condition].

7.4.2 *The robot's responses were unclear*

Other participants remarked that it was difficult to know if they were understood based on the ambiguous and seemingly inconsistent movements the robot made. "The movements [of the robot] were interesting. There were times when he moved his head. But I never knew if it was in response to what I'd said, or if it was just him moving as a movement in general. I wasn't really too sure. So again, if you compare it to an automated machine, you never know for sure if you've been fully understood or not" [Partial mimic condition]

Also, just as we found in our pilot study, in this study a few participants said that they wanted the robot to make non-speech sounds to indicate understanding or that it was paying attention to them. (For example, "Mm-hmm".)

7.4.3 *Facial expression frequency*

Two participants in the non-mimic condition made comments about the amount of facial expressions the robot made. One participant said they expected the robot to make more facial expressions, though they weren't sure why they had such an expectation. And another said, "It was quite difficult to interact with [Virgil] because I think that his facial expressions were a bit difficult to pick up. Like how he's responding to what I'm saying. His eye movements were there, but I think if there were some facial movements as well that would be more helpful for interacting with him."

No participants in the other groups made comments regarding facial expression frequency.

7.4.4 *Machine-like movements*

A few participants in the Full and Partial Mimic conditions said that the head movements were too erratic or jerky. For example, "I guess it didn't seemed like he moved much for awhile. I think maybe at the end he started to nod. But yeah, it seemed quiet mechanical. It was like somebody just flipped the switch and he decided to start moving, nodding to agree with me or whatever" [Partial mimic condition].

7.4.5 *Response appropriateness*

Three participants in the full mimic condition remarked that the robot made appropriate responses to what they were saying by moving its head and nodding in agreement. "[The robot was an amicable conversation partner] because you are still getting a sense that it's listening and paying attention to what you said." However, a fourth participant in this condition said that the robot's movements were *not* in tune with the conversation. But this may be explained by individual differences between participants—more data are needed to draw any strong conclusions.

Though it is worth noting that no participants in the other two conditions commented directly on the appropriateness of the robot's response, so it seems we can find some support for the idea that full mimicking may be worth pursuing if one wishes to make a conversational robot that is capable of rapport-building.

7.4.6 *Candor*

Two participants, both in the partial mimic condition, said they felt uncomfortable describing anything personal to the robot. One attributed this to the jerky, non-humanlike movements of the robot. The other said it was because there was no verbal feedback from the robot, "I would have talked more probably to a human about these things. But it's very hard to talk to something that's not talking back."

7.4.7 *Trust*

One participant in the partial mimic condition (who also mentioned talking to the robot was like talking to an automated machine) expressed a lack of confidence in the robot's ability to hold its end of the conversation. "I have the same impression of an automated machine as I do of this. I have no confidence in automated machines in that they don't work in the same way as a human."

7.4.8 *Robot feelings*

A few participants remarked on response appropriateness in response to them (as discussed in Sect. 7.4.5); however, only one remarked on the robot's emotions. This participant, in the non-mimic condition, said, "It was difficult to pick up what [Virgil's feelings were]. Oh, I guess the robot doesn't feel, actually! *laughs.* But it was hard to pick up what its expressions were." In spite of the self-correction, this participant later went on to say, "I found it difficult to know [Virgil's] feelings." It is unclear if these statements were due to the fact that this participant was not a native English speaker or if the participant was truly attributing emotions to the robot. Nonetheless, it is an interesting finding from an empathy perspective.

7.4.9 *Positive vs. negative emotions*

One participant in the partial mimic condition raised an interesting idea with regard to positive vs. negative emotions.

During the course of this participant's interaction, only positive things were discussed. The participant said, "I think it's easy to respond to positive emotions because you don't really need much of a response. You probably need more of a response to [the] negative [emotions other people express]. If you had to say something sad, or a bad memory, you'd want [the robot] to look sympathetic back. Whereas if you're telling a happy memory, you don't really care what the response is because it's a happy memory anyway and you already feel good. [For a positive emotion] the [other] person doesn't need to do anything back. Even if they look [at you] blankly it doesn't really matter. Whereas if you tell someone something sad and they look blank, it's not really good."

This was a very helpful response, because we did not specifically ask participants to recall positive or negative memories—just their first memories of Cambridge. In the future we may use more formal emotion elicitation to explore this area further.

7.5 Discussion

From this preliminary study, we have learned several lessons about enabling human-robot rapport via head mimicry. First, it seems that we may need to develop new metrics for evaluating affective interactions between humans and robots, as the self-report Interactant Satisfaction measure we chose for this experiment did not provide us with any sort of significant result. While this may be due to the number of participants we had in our experiment, it may also be that some of the questions are ill-suited toward interacting with a machine; particularly one that is silent. As Weiss et al. have found [23], it is difficult to use questionnaires to assess user-experience in general for human-robot interaction, it seems this may also be true for affective human-robot interaction. Thus, more follow-on work is needed to address this issue.

Gestural analysis of our data revealed a number of interesting features. First, we found a significant difference, with a large effect size, between the amount of gesturing men and women performed when interacting with the robot. This can probably be explained by the fact the men overall make more postural shifts than women in dyadic human-human interaction [4]; however, it's interesting to note that this effect was also seen with dyadic human-robot interaction.

Our gestural analysis also revealed that our participants often didn't move in the way we expected they would. To solely rely on head gestures to dictate our robot's movement may not have been sufficient for participants in the partial and non-mimicking conditions. Also, some participants slouch and some lean forward—this also could affect our ability to accurately track their faces in real-time.

A third revelation from our gestural analysis was the idea of coordinated gestures, such as co-nodding. This finding was a surprise to us, and something we plan to explore in more detail in the future. It may be the case that such coordination leads to more positive interaction with robots, as is this is definitely the case in human-human interaction [5].

Participants provided a wide variety of helpful feedback during the qualitative interviews that we conducted at the end of the experiment. In particular, the ideas of response appropriateness, response clarity, and positive vs. negative emotions will be very helpful as we embark on future research on enabling human-robot rapport.

Acknowledgements The authors would like to thank Shazia Afzal, Daniel Bernhardt, Ian Davies, Martin McDonnell, Simon Moore, and Graham Titmus for their assistance.

References

1. Arduino: Arduino physical computing platform (2009). http://www.arduino.cc. Last accessed 15 Apr 2009
2. Bailenson JN, Yee N (2005) Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. Psychol Sci 16:814–819
3. Davis MH (2006) Empathy. In: Stets JE, Turner JH (eds) Handbook of the sociology of emotions. Springer, New York
4. Frances SJ (1979) Sex differences in nonverbal behavior. Sex Roles 5(4):519–535
5. Gill S (2004), Body moves and tacit knowing. Cognition and technology: co-existence, convergence, and co-evolution, p 241
6. Gratch J, Wang N, Gerten J, Fast E, Duffy R (2007) Creating rapport with virtual agents. In: IVA '07: Proceedings of the 7th international conference on intelligent virtual agents. Springer, Berlin, pp 125–138
7. Harrison C, Hudson SE (2008) Pseudo-3d video conferencing with a generic webcam. In: International symposium on multimedia. IEEE Press, New York, pp 236–241. doi:ieeecomputersociety.org/10.1109/ISM.2008.12
8. Hatfield E, Cacioppo J, Rapson R (1992) Emotional contagion. In: Clark MS (ed) Review of personality and social psychology: emotion and social behavior. Sage, Thousand Oaks, pp 151–171
9. Hatfield E, Cacioppo J, Rapson R (1994) Emotional contagion. Cambridge University Press, Cambridge
10. Kanda T, Kamasima M, Imai M, Ono T, Sakamoto D, Ishiguro H, Anzai Y (2007) A humanoid robot that pretends to listen to route guidance from a human. Auton Robots 22(1):87–100
11. Kang SH, Watt JH, Ala SK (2008) Communicators' perceptions of social presence as a function of avatar realism in small display mobile communication devices. In: HICSS '08: proceedings of the proceedings of the 41st annual Hawaii international conference on system sciences. IEEE Comput Soc, Los Alamitos, p 147. doi:10.1109/hicss.2008.95
12. Kang SH, Watt JH, Ala SK (2008) Social copresence in anonymous social interactions using a mobile video telephone. In: CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on human factors in computing systems, Florence, Italy. ACM, New York, pp 1535–1544. doi:10.1145/1357054.1357295
13. Nadel J, Simon M, Canet P, Soussignan R, Blancard P, Canamero L, Gaussier P (2006) Human responses to an expressive robot. In: Proceedings of the sixth international workshop on epigenetic robotics, Lund University Cognitive Studies, pp 79–86
14. Oberman LM, Winkielman P, Ramachandran VS (2007) Face to face: Blocking facial mimicry can selectively impair recognition of emotional expressions. Soc Neurosci 2(3):167–178

15. Reas C, Fry B (2003) Processing: a learning environment for creating interactive web graphics. In: SIGGRAPH '03: ACM SIGGRAPH 2003 Web Graphics. ACM, New York, p 1. doi:10.1145/965333.965356
16. Riek LD (2007) Realizing hinokio: candidate requirements for physical avatar systems. In: HRI '07: proceedings of the ACM/IEEE international conference on human-robot interaction. ACM, New York, pp 303–308. doi:10.1145/1228716.1228757
17. Riek LD, Afzal S, Robinson P (2009) Do affect-sensitive machines influence user behavior? In: Proceedings of the AISB symposium on the social understanding of artificial intelligence (SSoAI). Edinburgh, UK
18. Riek LD, Robinson P (2008) Real-time empathy: facial mimicry on a robot. In: Workshop on affective interaction in natural environments (AFFINE) at the international ACM conference on multimodal interfaces (ICMI 08). ACM, New York
19. Ross DM, Menzler S, Zimmermann E (2008) Rapid facial mimicry in orangutan play. Biol Lett 4(1):27–30 doi:10.1098/rsbl.2007.0535
20. Sidner CL, Lee C, Morency LP, Forlines C (2006) The effect of head-nod recognition in human-robot conversation. In: HRI '06: proceedings of the 1st ACM SIGCHI/SIGART conference on human-robot interaction. ACM, New York, pp 290–296. doi:10.1145/1121241.1121291
21. Steiner HC The firmata protocol (2008). http://www.firmata.org. Last accessed 15 Apr 2009
22. Walters ML (2008) The design space for robot appearance and behaviour for social robot companions. Ph.D. thesis, University of Hertfordshire. http://hdl.handle.net/2299/1806
23. Weiss A, Bernhaupt R, Tscheligi M, Yoshida E (2009) Addressing user experience and societal impact in a user study with a humanoid robot. In: Proceedings of the AISB symposium on new frontiers in human-robot interaction. Edinburgh, UK
24. Wilson PI, Fernandez J (2006) Facial feature detection using Haar classifiers. J Comput Small Coll 21(4):127–133
25. WowWee: WowWee alive chimpanzee robot (2008). http://www.wowwee.com. Last accessed 15 Apr 2009