

generally selected adherence to the collaboration plan to be more relevant.

Imagine you have made the peanut butter sandwich and passed it to Mary to cut it in half. Which of the following two actions is **more relevant**?

A. Mary starts crying since she cut her finger with a knife.

B. You begin to boil the water to boil the eggs for your second sandwich.

C. Equally relevant.

Figure 5.9: Example relevance question.

5.1.4 Discussion

As shown in the results tables in sections 5.1.3 through 5.1.3, the human participants agreed 100% on some questions, while on some other questions there was a much lower level of agreement. Our results indicate that people largely performed as our hypothesis predicted. The p -values obtained based on a binomial distribution show the probability of human participants' answers being generated from a random set. The very small p -values indicate that the data set is not random; in fact, the high percentage of similarity confirms our hypothesis and shows that the algorithms can help us to model appraisal in a collaboration. The very low level of agreement on a handful of questions may indicate algorithm components that require further refinement before implementation.

5.2 End-to-End System Evaluation

As mentioned earlier, collaborative robots need to take into account humans' internal states while making decisions during collaboration. Humans express emotions to reveal their internal states in social contexts including collaboration [35]. Due to the existence of such expressions robots' emotional-awareness can improve the quality of collaboration in terms of humans' perception of performance and preferences. Hence, collaborative robots need to include affect-driven mechanisms in

their decision-making processes to be able to interpret and generate appropriate responses and behaviors. Our aim in this setup was to study the importance of emotional awareness and the underlying affect-driven processes in human-robot collaboration. We examined how emotional-awareness impacts different aspects of humans' preferences by comparing the results from our participants collaborating with an emotion-aware and an emotion-ignorant robot.

5.2.1 Implementation

The implementation of this user-study included three separate parts. The first part incorporated the Affective Motivational Collaboration Framework consisting of all Mental Processes (see left-side of Figure 5.10) as we described in Chapter ???. The second part was implemented to receive action commands from the framework and forward them to the robot to control joints and actuators (see right-side of Figure 5.10). A wizard was the third part of this setting. The wizard did nothing but inform the robot/framework whether the current task performed by either the robot or the participant was achieved successfully. The wizard was completely invisible to the participants, and the wizard had no impact on the robot's decision other than providing input regarding tasks' failure or success.

Framework

The framework includes all of the mechanisms depicted as mental processes in Figure 5.10 along with the mental states. The mental states shown in Figure 5.10 comprise the knowledge base required for all of the mechanisms in the overall model. The details about these mental processes and mental states are described in Chapters 3 and ??. In this user-study, the Collaboration mechanism uses a hierarchy of goals associated with tasks in the hierarchical task network structure depicted in Figure 5.11.

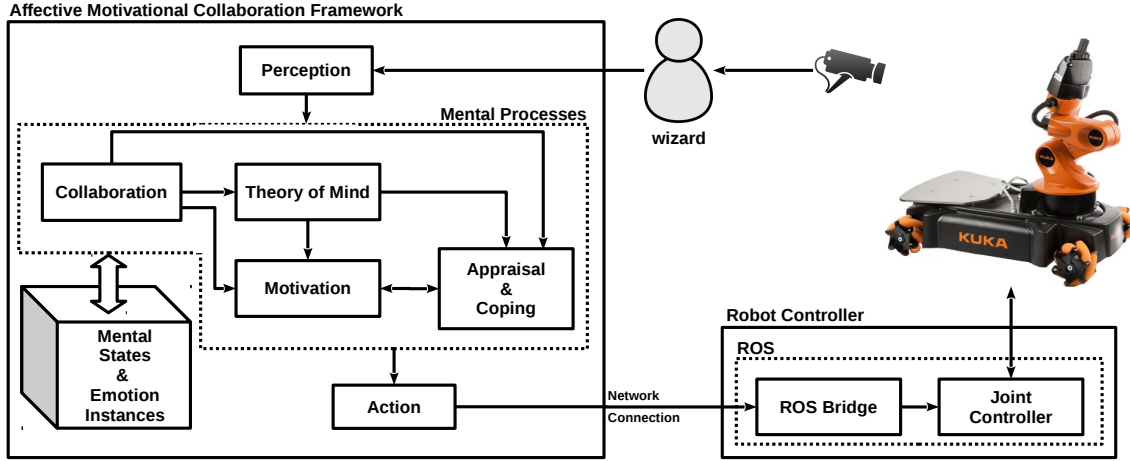


Figure 5.10: Computational framework based on Affective Motivational Collaboration theory (arrows indicate primary influences between mechanisms and data flow).

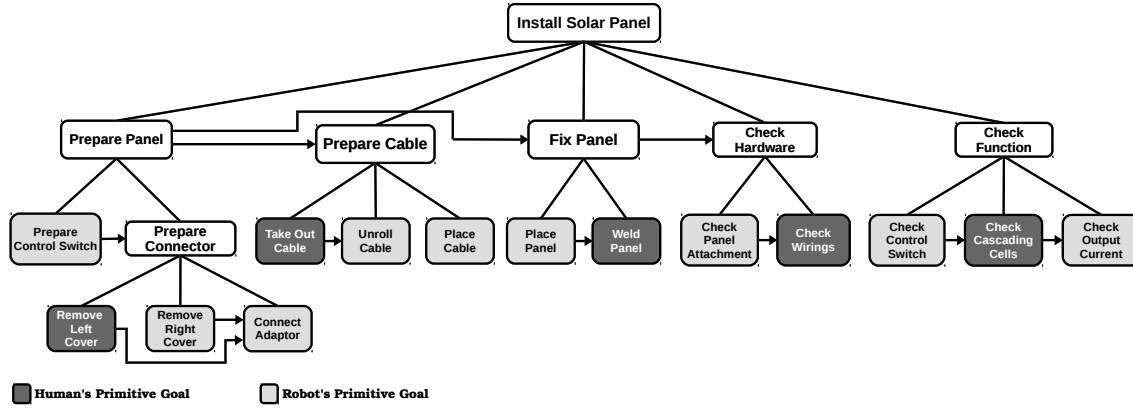


Figure 5.11: Collaboration structure used as the task model.

Robot Controller

The robot controller is comprised of two major components: 1) ROS-bridge and 2) joint controller (see Figure 5.10). ROS-bridge¹ provides an API to ROS functionality for non-ROS programs which enables us to send action commands from our framework (implemented in JAVA) to the robot's joint controller. The joint controller receives action commands and translates them into actual joint and actuator commands and sends them to the robot.

¹http://wiki.ros.org/rosbridge_suite

5.2.2 Experimental Scenario

Our scenario was based on a table top turn-taking game that we designed to simulate the installation of a solar panel. Participants collaborated one-on-one with our robot to complete all the given tasks required to install the solar panel. All of the tasks consisted of picking up and placing collaborators' available pegs on predefined spots on the board (see Figure 5.12). Each pick-and-place was associated with the robot's or the participant's task. The robot and the participants had their own unique primitive tasks that they had to accomplish in their own turns. The final goal of installing a solar panel required the robot and the participants to accomplish their own individual tasks. Failure of any task could create an impasse during the collaboration.

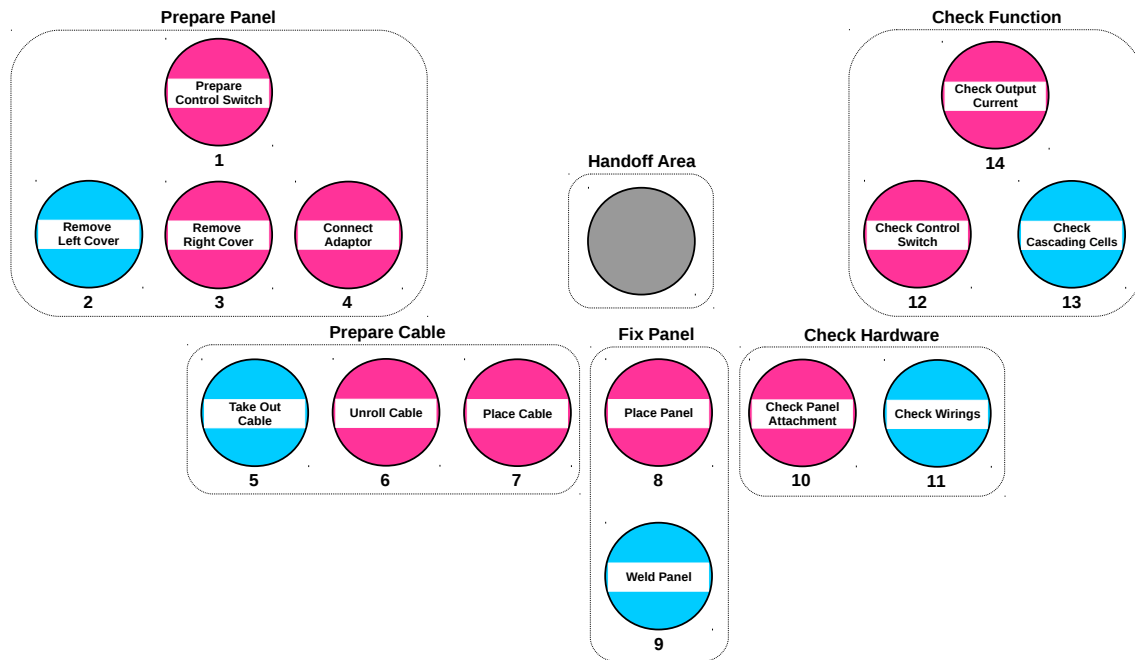


Figure 5.12: The layout of the available spots for the human and the robot to place their pegs during the collaboration.

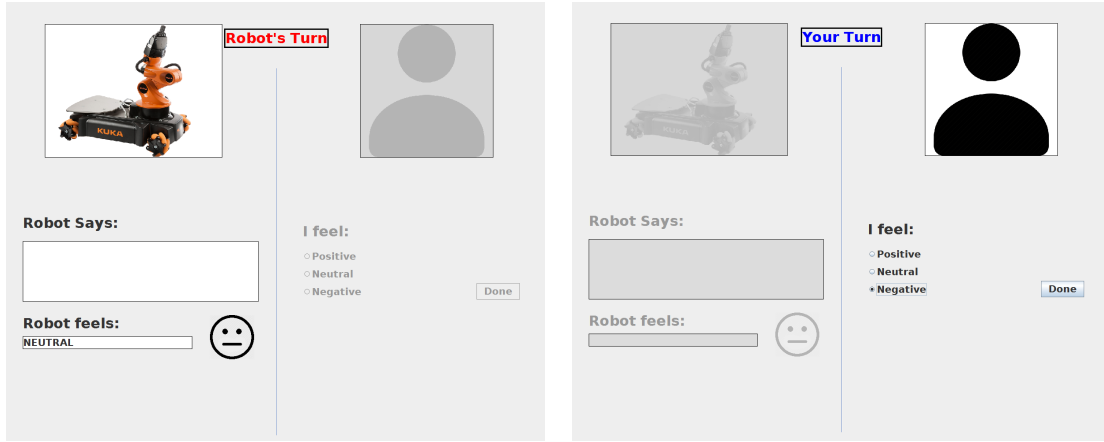


Figure 5.13: The Graphical User Interface (GUI) used during interaction.

The Robot

We conducted our experiment based on a KUKA Youbot (see Figure 5.14). The robot was stationary on top of a desk and was able to pick up and place available pegs corresponding to the robot’s task. The robot was operated based on Robot Operating System (ROS – indigo) and was receiving commands through the ROS-bridge from our Affective Motivational Collaboration framework (see Figure 5.10). We provided a simple GUI using a touch-screen monitor (see Figure 5.13 and) to a) express the robot’s positive, negative or neutral emotion through an emoticon, b) display the robot’s utterances, c) control turn-taking process of the collaboration, and d) let the participants express (report) their positive, negative or neutral emotion for each turn. The robot used MaryTTS an open-source, multilingual Text-to-Speech Synthesis platform to provide corresponding speech for its utterances in English.

Interaction Paradigms

At the beginning of each collaboration the robot asked each participant to achieve the overall shared goal, i.e., “installing the solar panel”. Then, before working towards a new goal, the robot informed the participant about the higher level non-

primitive goal (e.g. Prepare Panel – see Figure 5.11) of which the primitives were going to be working towards. The same procedure was used by the robot if there was a decision to switch to another nonprimitive due to the failure of a task in achieving the current goal. After achieving a new primitive goal, the robot either informed the human that it would pursue the next goal, or it informed and passed the turn to the human to execute the next task with respect to the human’s goal. In case of the human’s turn, the robot waited for the human to do a task, then the wizard let the robot know whether the human’s goal was achieved or not. Afterwards the robot made a decision about which goal to pursue and informed the human accordingly. The same procedure was applied to both conditions.

The robot interacted via a) speech, b) the corresponding utterance on the screen, c) negative, positive and neutral expression of emotion through an emoticon on the screen. There were two conditions of the robot: 1) emotion-aware and 2) emotion ignorant. The robot used neutral expression in the case of emotion-ignorance. The interaction was controlled autonomously by the framework we discussed in Section 5.2.1 in both the emotion-ignorant and the emotion-aware cases. The reasoning about which task should be done and controlling the robot was entirely autonomous. Only the perception of the task failure or achievement by the robot or by the participant was done by a wizard monitoring the collaboration outside of the test area. The interaction was structured based on the exact same goals in the same HTN for both conditions. The robot was using the same utterances in both conditions. In the emotion-aware condition the robot used a different behavior in comparison with the emotion-ignorant condition only if the participant was expressing a negative emotion in the event of a failure; i.e., the robot’s utterances were identical in emotion-ignorant and emotion-aware cases if in the latter the participant reported (expressed) a positive or a neutral emotion.

Three different behaviors could be generated only in the emotion-aware condition. These three behaviors were 1) mitigating the human’s negative emotion and postponing its own task to help the human, 2) goal-management to switch to another

goal which has lower cost with respect to the human’s negative emotion, and 3) task delegation to the human to overcome the impasse. In each run, the human had two pre-coordinated task failures, and the robot had one. If the human expressed negative emotion after the first human task-failure, the robot responded by mitigating the human’s negative emotion by saying “It was not your fault. I can help you with this task” and helping the human by providing a peg to fulfill the human’s task. If the human expressed negative emotion after the second human task-failure, the robot informed the human that they could proceed with another task to save time while simultaneously requesting a new peg (i.e. help) from the supervisor. If the human expressed negative emotion as a result of the robot’s task failure, the robot requested help from the human (who had the correct peg). In the event that the human expressed positive or neutral emotion during these three failures, the robot behaved identically in the emotion-ignorant and the emotion-aware cases, by asking the supervisor for help.

Environment and Tasks

The environment was set up in the Human-Robot Interaction lab and included the robot, the collaboration board on top of a desk, and the participant standing in front of the robot on the other side of the board (see Figure 5.14). One of the experimenters monitored the interactions using a live stream of a camera in a different room. The experimenter provided only the required perception, i.e., decision on success or failure of the tasks for the robot, through the entire time of the collaboration (see Section 5.2.2).

The tasks were defined based on the HTN structure shown in Figure 5.11 and were executed in a turn-taking fashion by either of the collaborators. For each task either the robot or the participant was responsible to pick up one of the corresponding pegs from their own inventory and place it on the right spot which was colored and tagged the same as the associated peg. Some pegs and corresponding spots on the board had hidden magnets which prevented the pegs from standing upright.

Any peg that fell over was considered a failed task.

5.2.3 Hypotheses and Methodology

Hypothesis

The non/social functions of emotions impact a collaboration process. Human collaborators prefer to collaborate with others whose behaviors are influenced by these functions of emotions depending on the context. We developed seven hypotheses on positive influence of emotion-awareness and usefulness of emotion function during collaboration:

Hypothesis 1. Participants will feel closer to the emotion-aware robot rather than the emotion-ignorant robot.

Hypothesis 2. Participants will find the emotion-aware robot to be more trustworthy than the emotion-ignorant robot.

Hypothesis 3. Participants will find the emotion-aware robot to have better performance in collaboration than the emotion-ignorant robot.

Hypothesis 4. Participants will find the emotion-aware robot to be more understanding of their feelings than the emotion-ignorant robot.

Hypothesis 5. Participants will find the emotion-aware robot to be more understanding of their goals than the emotion-ignorant robot.

Hypothesis 6. Participants will feel more satisfied about the collaboration when working with the emotion-aware robot rather than emotion-ignorant robot.

Hypothesis 7. Participants will perceive higher level of mutual satisfaction with the emotion-aware robot than emotion-ignorant robot.

Procedure

Participants were first given a brief description of the purpose of the experiment. After the short introduction, they were asked to review and sign a consent form. Participants were then provided with a written instruction of their task and the rules

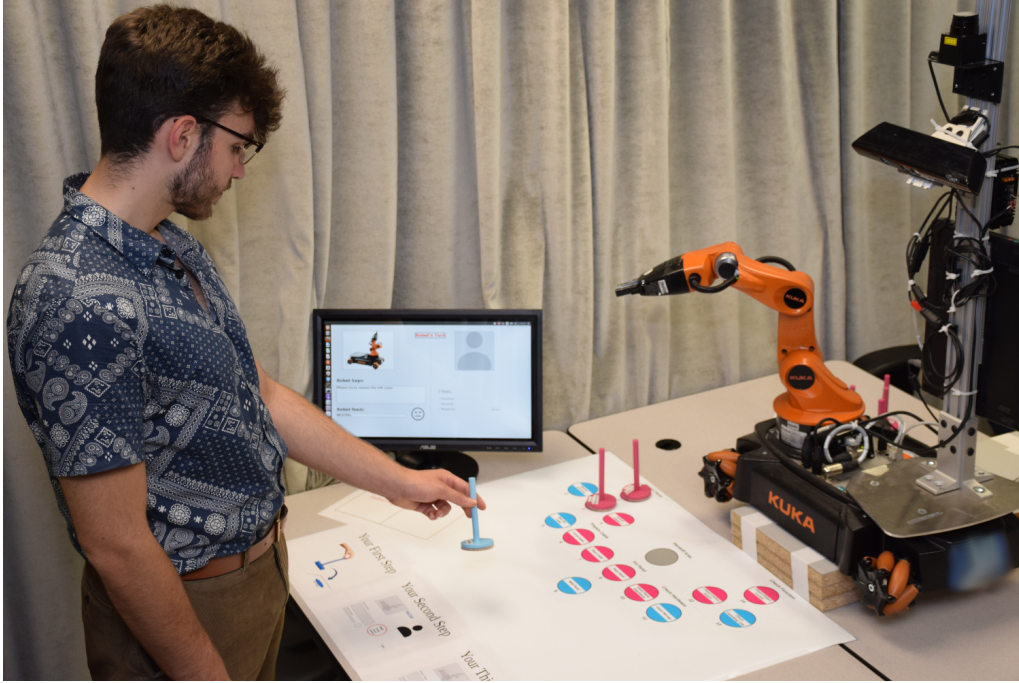


Figure 5.14: Experimental setup.

for collaborating with the robot. Then, one of the experimenters lead them into the experiment room and asked the participants to answer pre-experiment questionnaires. Afterwards, the experimenter went through all the details of the instructions with the participants standing in front of the collaboration board and the robot. The experimenter confirmed participants' correct understanding of the tasks and informed them of types of task failures that might occur during the collaboration. Participants were told that researchers were developing a collaborative robot and would like their help in evaluating their design. Participants were provided with identical instructions and randomly assigned to complete either the emotion-aware or the emotion ignorant condition first. They were told that, after their collaboration with the robot, they would be asked to answer a questionnaire on their experience. After completing the first round of collaboration, participants answered a post-experiment questionnaire that measured their perceptions of the robot, the task, and the collaboration procedure. After answering the first post-experiment questionnaire, participants were told that they were going to collaborate with the

robot one more time and the robot might not necessarily have the same collaborative behavior. After completing the second round of collaboration, participants were asked to answer the second post-experiment questionnaire which consisted of the same questions as the first post-experiment questionnaire. After all, participants were asked to answer an open-ended questionnaire which measured their perception of difference between two runs, their preference of collaborative robot between two runs, and their reasons of preference.

Measurements

In our study two basic conditions of the robot were tested: a) the emotion-ignorant condition, b) the emotion-aware condition. We measured participants' recall of the collaborative behaviors presented by the robot using an open-ended post-experiment questionnaire. We also specifically asked the participants what behavior of the robot they liked during their collaboration. We also evaluated participants' levels of satisfaction, trust, goal achievement, mutual understanding of goals, mutual understanding of feelings, mutual agreement, and also participants' beliefs about the efficiency of collaboration and their feeling of robot's collaborative behaviors. Seven-point Likert scales were used in these questionnaire items.

Participants

A total of 37 participants participated in the experiment in 74 trials. Participants were recruited from Worcester Polytechnic Institute's students and staffs as well as other civilians recruited from outside of the campus. The ages of the participants varied between 19 and 74 with an average of 34.2 years before our screening of 4 participants based on our sanity check questions. After this screening the ages of the participants varied between 19 and 54 with an average of 30.8 years old. Of the 33 participants, 21 were female and 12 were male. Each participant participated in 2 trials. In one trial the robot was aware of human's emotion and in the second

trial the robot was ignoring human’s emotion. The order of these two trials were randomly assigned to each participant. In general we used emotion-ignorant robot first in 16 experiments, and emotion-aware robot first in 17 experiments.

5.2.4 Results

As discussed in Section 5.2.3, results of the user study were gathered through a 31-question Likert-scale survey that was given to each participant after each run with the robot, and through a 5-question open-ended summary questionnaire at the end of the experiment.

7-Point Likert Scale Survey Results

As mentioned previously, the 7-point Likert scale survey was administered at the end of the emotion-ignorant run and at the end of the emotion-aware run for each participant. The 31 questions are generally categorized to evaluate the humans’ perceptions of the following seven categories, with 3-7 questions per group: (1) the likability of the robot (2) the level of trust the human feels in the robot (3) the human’s perception of the robot’s performance (4) the human’s perception of the robot’s understanding of human’s emotions (5) the human’s perception of the robot’s understanding of human’s and collaboration’s goals and objectives (6) the human’s feeling about the collaboration and (7) the human’s perception of the human’s and robot’s mutual satisfaction with each other as collaborative partners. The questions presented are provided in Figure 5.15.

Question Category	Question	Question Number
Likability	I felt close to the robot.	Q1
	I would like to continue working with the robot.	Q2
	I like the robot.	Q3
	The robot was interesting.	Q4
Trust	I trust the robot.	Q5
	It was easy to express myself to the robot.	Q6
	I trust the robot to perform appropriately in our collaboration.	Q7
	I am confident in the robot's ability to help me.	Q8
	I trust the robot to assess my feelings appropriately in our collaboration.	Q9
Robot's Performance	The robot was repetitive.	Q10
	The robot made efficient decisions.	Q11
	The robot's decisions improved my performance during the collaboration.	Q12
Robot's Understanding of Human's Emotions	The robot understood my emotions.	Q13
	The robot is sometimes confused about what I feel about our activities.	Q14
	I feel that the robot, in its own unique ways, is genuinely concerned about me.	Q15
	The robot understands some of my feelings and takes them into account in our collaboration.	Q16
	The robot does not understand how I feel during our collaboration.	Q17
Robot's Understanding of Goals	The robot does not understand what we are trying to accomplish.	Q18
	The robot does not understand what I am trying to accomplish.	Q19
	The robot perceives accurately what my objectives are.	Q20
	The robot was committed to the collaboration.	Q21
Human Feeling about Collaboration	I find what the robot and I are doing is unrelated to my goals.	Q22
	I find what I am doing with the robot confusing.	Q23
	The robot and I are working towards mutually agreed-upon goals.	Q24
	The robot and I collaborate on setting goals for us to work on.	Q25
	The robot and I agree on what is important for us to work on.	Q26
	I believe that the robot and I achieved the goals we set.	Q27
	I am satisfied with the outcome of our collaboration.	Q28
Satisfaction of Collaborative Partner	The robot was satisfied with my collaborative behavior.	Q29
	I was satisfied with the robot.	Q30
	I understand the robot, and I think it understands me, at least in the best way it can.	Q31

Figure 5.15: The 31 Likert scale questions organized according to their groups.

The results were analyzed using a two-tailed paired t-test to analyze the difference of means between the emotion ignorant and the emotion-aware condition. Refer to Figures 5.18 - 5.22 for the results. As mentioned in Section 5.2.3, participants were randomly assigned to complete either the emotion-ignorant or the emotion-aware run first; analysis of the results revealed no statistically significant difference or consistent pattern based on which run the participant completed first.

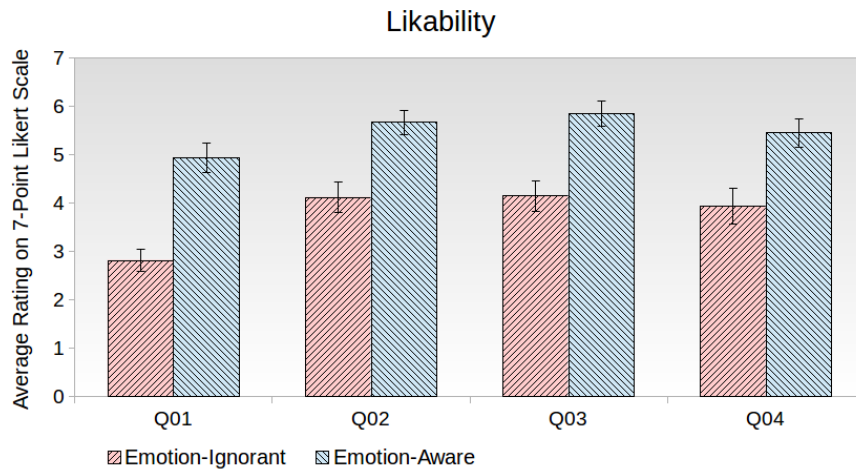


Figure 5.16: Results of the Likert scale survey for Likability questions. The p-value for the difference between means is $\ll 0.001$ for all questions.

Likability of the Robot

Questions 1 through 4 addressed the likability of the robot. As shown in Figure 5.16, participants rated the emotion-aware robot 1.5-2.1 points higher than the emotion-ignorant robot. These results indicate that participants felt closer with and preferred working with the emotion-aware robot; these results support Hypothesis 1, which stated that humans would prefer to work with the emotion-aware robot over the emotion-ignorant robot.

Human Trust in the Robot

Questions 5-9 were designed to measure the degree of trust that the human participants felt in the robot. As shown in Figure 5.17, participants trusted the emotion-

aware robot, on average, a minimum of 1.4 points more than the emotion-ignorant robot, both in general and in terms of collaboration performance. In Question 5, participants rated a general statement of trust 1.5 points higher in the emotion-aware case. Additionally, in Question 7, participants rated their trust in the emotion-aware robot to perform appropriately during collaboration an average of 5.9 on a 7-point Likert scale, where 7.0 would indicate maximum trust; this indicates an acceptable level of trust in the robot’s collaborative abilities. These results support Hypothesis 2, that posits that human participants would find the emotion-aware robot to be more trustworthy than the emotion-ignorant robot.

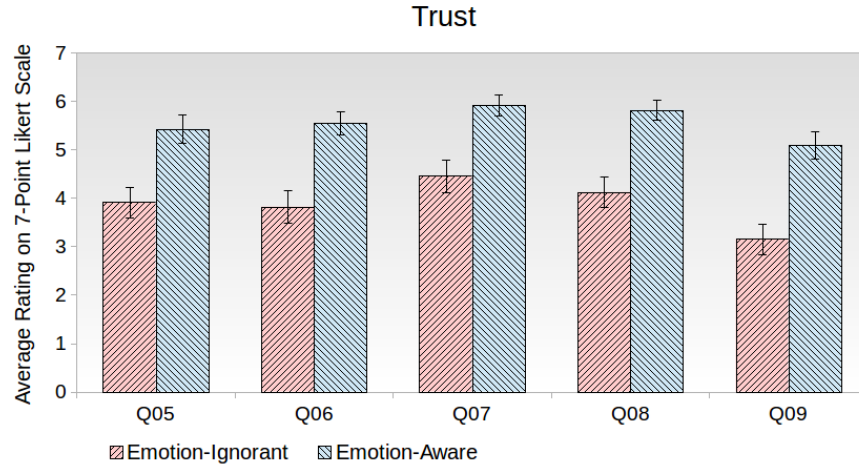


Figure 5.17: Results of the Likert scale survey for questions related to trust. The p-value for the difference between means is $\ll 0.001$ for all questions.

Perception of the Robot’s Performance

Question 10 (which is reverse-scored) measures the participant’s perception of repetitiveness in the robot during the collaboration. In both conditions, participants rated the robot as moderately repetitive, with the emotion-ignorant robot’s average response being about 1.1 points higher than the emotion-aware. This result correlates with several of the open-ended responses which described the emotion-aware robot’s behaviors as “cute” and “interesting”, refer to Section 5.2.4. Question 11, which asks

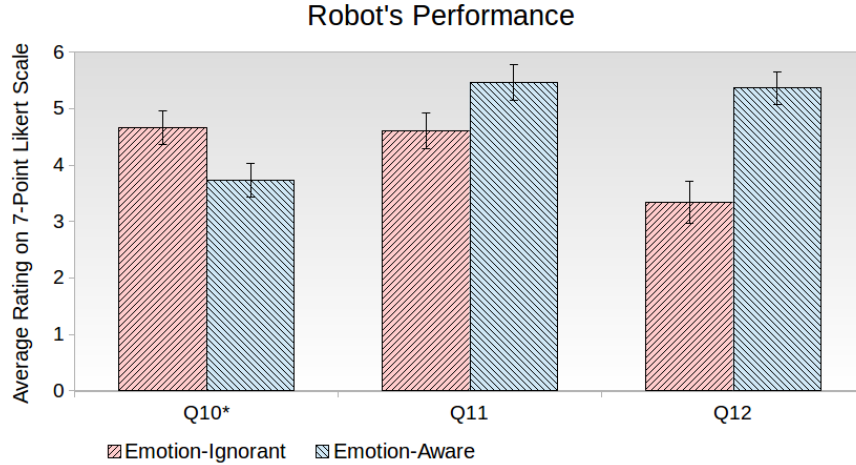


Figure 5.18: Results of the Likert scale survey for questions related to the robot's performance. The p-value for the difference between the means for questions 10, 11 and 12 are 0.001, 0.063 and $\ll 0.001$, respectively.

about the efficiency of the robot's decisions is the only question of the 31 questions that did not have a statistically significant difference between the emotion-aware and the emotion-ignorant case. This correlates with the result of the open-ended question asking which condition of the robot exhibited behaviors that could prevent human error (refer to 5.2.4); in response to this question, several respondents stated that it may be quicker or simpler to call the supervisor in the event of a task failure, rather than changing the order of the tasks. According to the results from Question 12, the participants felt that the emotion-aware robot's decisions during collaboration improved their own performance, with an average rating of 5.4, while the emotion-ignorant robot only received an average rating of 3.3, indicating that participants felt it was not able to interact in such a way as to increase the human's performance; refer to results from Question 6. These results support Hypothesis 3, which posited that humans will perceive the emotion-aware robot as being more capable than the emotion-ignorant robot.

Robot's Understanding of Human Emotions

In Questions 13 through 17, participants evaluate the robot's understanding of hu-

mans' emotions. In questions 13, 15, and 16, participants rated the emotion-aware robot, on average, a minimum of 1.8 points higher than the emotion-ignorant robot. In response to questions 14 and 17, which are reverse-scored, participants ranked the emotion-ignorant robot 1.2 and 2.0 points higher, respectively, than the emotion-aware robot. The results of all five questions in this category support Hypothesis 4.

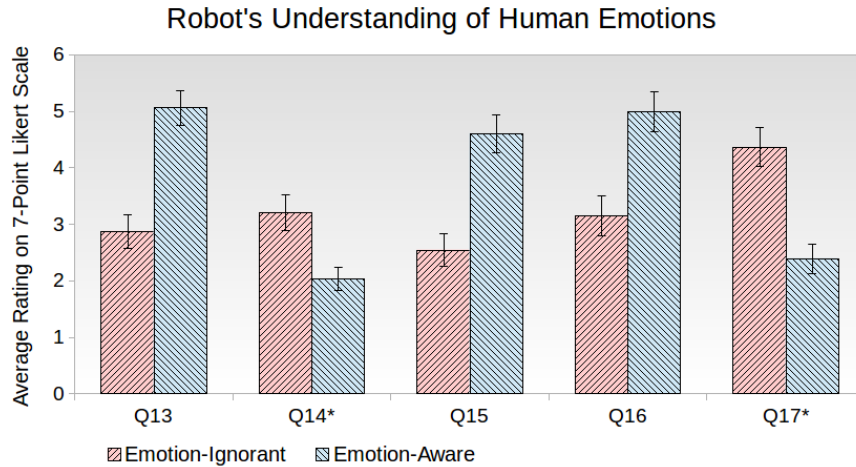


Figure 5.19: Results of the Likert scale survey for the questions related to the robot's understanding of human emotions. The p-value for the difference between the means is $\ll 0.001$ for all of the questions except Question 14, for which the p-value is 0.003.

Robot's Understanding of Human and Collaboration Goals

Questions 18 and 19 were reverse-scored questions intended to determine whether the humans felt that the robot understood the shared collaboration goal and the human's personal goal, respectively. For both conditions of the robot, the average scores were lower than 3.5, indicating that the human's perceived the robot as having some understanding of the goals. For both questions, the emotion-ignorant robot's average score was significantly higher than the emotion-aware robot's score. Similarly, Question 20 was a measure of whether the human perceived that the robot correctly perceived the human's goal. On average, participants provided an average rating for the emotion-aware robot that was 1.5 points higher than that for the emotion-ignorant robot. Question 21 measured the human perception of the robot's

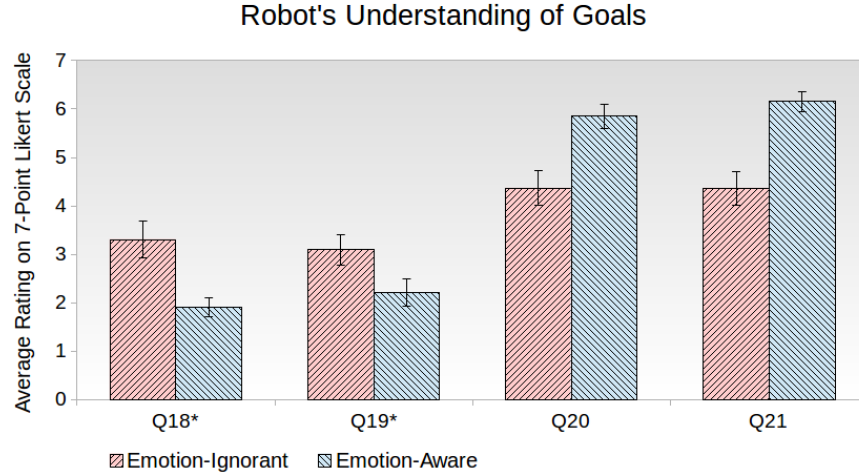


Figure 5.20: Results of the Likert scale survey for questions related to the robot’s understanding of goals. The p-value for the difference between the means for all questions is $\ll 0.001$, except Question 19, for which the p-value is 0.006.

commitment to the collaboration; for this measure, the average participant score assigned to the emotion-aware robot was 6.2 points out of a maximum of 7 points, indicating that the participants felt that the emotion-aware robot was strongly committed to the collaboration. The emotion-ignorant robot received an average rating of 4.4 points, indicating only moderate commitment. These results strongly support Hypothesis 5, which posits that humans will feel that the emotion-aware robot will better understand their goals than the emotion-ignorant robot.

Human’s Feeling about the Collaboration

Questions 22 through 28 were designed to gauge how the human participants felt about the partnership within the collaboration and the outcome of the collaboration. For each of the 7 questions, the participants ranked the emotion-aware robot as better than the emotion-ignorant robot, by a minimum, on average, of 0.8 points. Questions 24, 27 and 28 addressed whether the robot and the participant were working toward mutually agreed-upon goals and on the outcome of the collaboration; in the emotion-aware condition, participants rated the robot a minimum of 6.1 points,

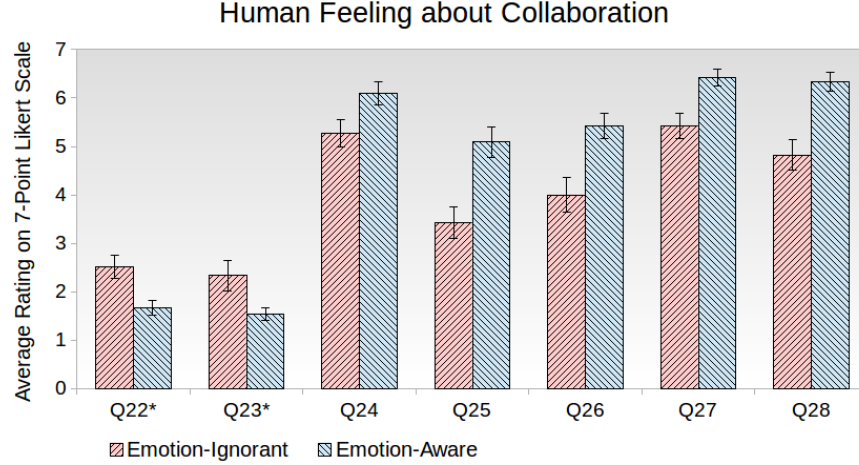


Figure 5.21: Results of the Likert scale survey for questions related to the human’s feeling about the collaboration. The p-value for the difference between the means is $\ll 0.001$ for questions 22, 25, 26, and 28. The p-value for Questions 23, 24 and 27 are 0.02, 0.008 and 0.001, respectively.

on average, while rating the emotion-ignorant robot 1-1.6 points lower, indicating that the participants felt a very strong sense of collaboration with the emotion-aware robot, and only a moderate sense of collaboration with the emotion-ignorant robot. Questions 25 and 26 address whether the robot and the participant set the collaboration goals together; these two questions have lower scores than Questions 24, 27 and 28, for both the emotion-aware and the emotion-ignorant case. The lower overall scores are likely due to the fact that the robot decides the task order or action in the event of failure in both conditions; however, the higher score in the emotion-aware case may indicate that emotional awareness can increase a feeling of collaboration. These results support Hypothesis 6 that humans will feel a greater sense of mutual collaboration and understanding about the collaboration with the emotion-aware robot.

Human Perception of Mutual Satisfaction with Collaborative Partner

Questions 29, 30 and 31 were designed to measure the human’s perception of the robot’s satisfaction with the human, the human’s satisfaction with the robot and

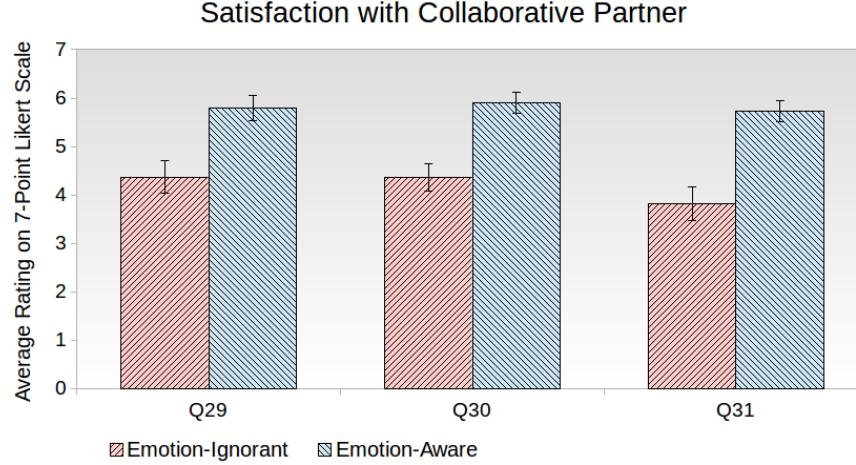


Figure 5.22: Results of the Likert scale survey for questions related to satisfaction with collaborative partner. The p-value for the difference between means is $\ll 0.001$ for all questions.

the mutual understanding between the human and the robot, respectively. The participants provided an average response in the emotion-aware condition of 5.8, 5.9 and 5.7 to Questions 29, 30 and 31, respectively, indicating a high level of mutual satisfaction; all three answers were about 1.4-1.9 points lower, on average, in the emotion-ignorant condition. These results indicate a higher level of satisfaction working with the robot in the emotion-aware condition, and strongly support Hypothesis 7, which posited that humans will feel a greater sense of mutual satisfaction with the emotion-aware robot than the emotion-ignorant robot.

Results from the Open-Ended Questionnaire

As described in Section 5.2.3, each participant answered an open-ended questionnaire at the end of the study. Figure 5.23 summarizes the questionnaire and which run users preferred for certain conditions (i.e. emotion-ignorant or emotion-aware). Note that some users chose not to state a preference regarding which run they preferred for certain conditions; because we were specifically interested in whether users preferred the emotion-aware case, we considered the ambiguous responses to be fail-

ures in the binomial analysis. The binomial analysis is based off of a population size of 33.

Question	Number of Participants Who Did Not Prefer One Run Over the Other *	Number of Participants Favoring Emotion-Aware Robot	p-value
Which of the two runs with the robot did you prefer?	0	33	0
In which of the two runs did the robot exhibit behavior that could be useful in a more complex task?	1	30	$\ll 0.001$
In which of two runs did the robot exhibit behavior that could prevent human error?	3	18	> 0.1
In which of the two runs did the robot exhibit behavior that could improve the efficiency of collaboration?	2	26	$\ll 0.001$
What was the most interesting behavior of the robot and in which run did it happen?	5	24	0.002

Figure 5.23: Open-ended questionnaire questions and results. (*Note: Because we are evaluating whether humans prefer an emotion-aware robot, these results are taken as negative test results when calculating the p-value using the binomial distribution. Only those participants who clearly indicated a preference for the emotion-aware robot are taken as positive test results.)

As shown in Figure 5.23, 100% of users unambiguously preferred the run with the emotion-aware robot. In general, this preference stemmed from a feeling of closeness and partnership, as seen in these responses: “the robot had emotions and responded to my emotions. Also, what it said about my failing was cute and aimed to make me feel better.” Another example is “I liked feeling needed and accounted for; I felt closer to the robot.” Finally, “I saw the changes in its feeling, which motivated me to care more about my act...I also liked that he asked me to correct its failure, although it could ask the supervisor.”

When asked in which of the two runs the robot exhibited behavior that could be useful in a more complex task, 90.9% chose the emotion-aware robot. In general, respondents thought that the emotion-aware robot was better at problem solving, more adaptable, and more capable of handling the social complexities that occur in collaboration, as shown in responses such as “The robot explained motives...which is important to keep a team communicating and on the same pace.” Also, “When we failed he initially switched to a new task and then came back to the originally failed task. It kept me from getting irritated and negative.” Finally, “The more

complex, the more necessary it is to understand how humans think and operate...an empathetic robot can adapt, encourage and help.” It is worth noting that one respondent preferred the emotion-ignorant case, saying “In a more complex task it might be better for the robot to take control and simply tell me what to do; trying to be understanding and collaborative wouldn’t be as important as doing the task correctly.”

The only question that did not provide statistically significant support in favor of the emotion-aware robot related to which case the robot exhibited behavior that could prevent human error. About 36.4% of respondents thought that the emotion-ignorant robot was more likely to prevent human error; however, all but one of these cited calling the supervisor as the main method of preventing human error, in spite of the fact that the instructions indicated that the robot’s need to call the supervisor counted against the collaboration. Of the 54.5% who thought that the emotion-aware robot was better at preventing human error, most cited the robot’s ability to console the human as the main behavior that could prevent human error. Respondents indicated that this enabled them to move on and feel better about the collaboration, as with this response: “The robot switched to a different task and we came back to an error later. This allowed my mind to move away from being frustrated. I was able to complete a different task which felt like a win - then come back and finish the error. Making my mind move away from frustration could definitely prevent more errors.”

When asked in which of the runs the robot exhibited behavior that could improve the efficiency of the collaboration, 78.8% responded with the emotion-aware case; of these, the vast majority stated that this was because of the robot’s ability to change the order of tasks in the event of a failure, and to ask the human for help.

Finally, when asked in which run the most interesting behavior occurred, 72.7% chose the emotion-aware condition. Of these respondents, 12 individuals stated that the robot’s attempt to console the human by saying “It was not your fault” in response to the human’s negative emotion that occurred as a consequence of the

human’s failed task was the most interesting behavior, and a majority mentioned that it actually made them feel more positive. Six participants referred to the robot’s ability to understand and express emotion. Several participants referred to the robot’s ability to communicate, including the ability to ask questions. Of those who responded with the emotion-ignorant case, most found the ability to call the supervisor, and mechanical functions, such as gripping, to be most interesting.

Impact of Demographics

As mentioned in Section 5.2.3, we recorded certain demographic information from each participant, including age and gender. We also had each participant complete several personality questionnaires. Although it was not the primary purpose of the study, we investigated the Likert scale results to determine if there were any relevant trends based on the demographics and personalities of the participants. A close study of the results did not reveal any identifiable pattern based on gender or personality.

Age did reveal an interesting pattern. We divided the participants into two groups, below 30 years of age and 30 or above. While question-by-question comparisons revealed only a few statistically significant differences based on age, a general pattern emerged. For all but four of the 31 questions presented, the younger age group reported higher scores than the older age group (or lower, in the case of reverse-scored questions) for the emotion-aware robot. In the emotion-ignorant case, the younger group tends to score the robot nearer to the same value as the older age group for all but seven questions, leading to a pattern in which the score drop between the emotion-aware and the emotion-ignorant case was more for the younger group than for the older group; the seven questions that broke this pattern were 7, 9, 11, 12, 18, 19 and 22.

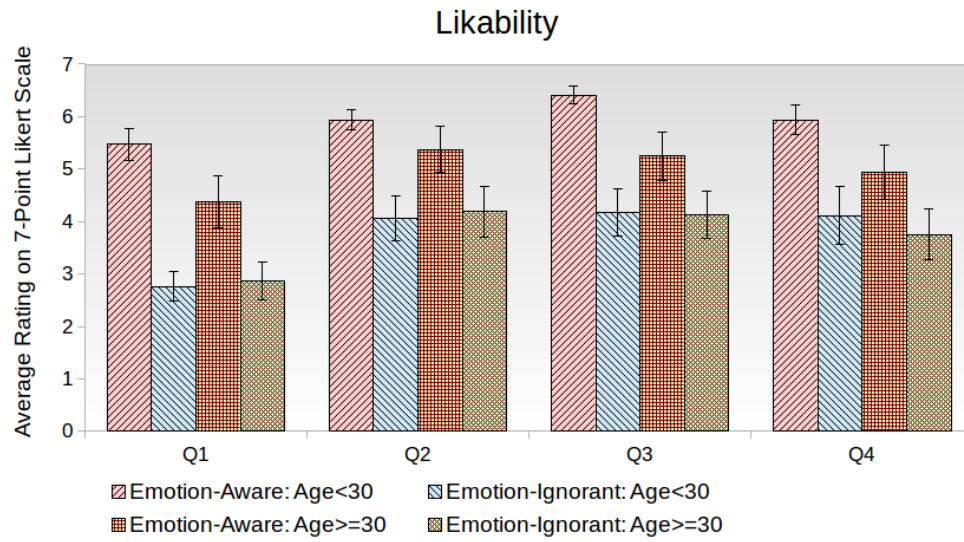


Figure 5.24: Impact of age on results of Likert scale questions related to likability.

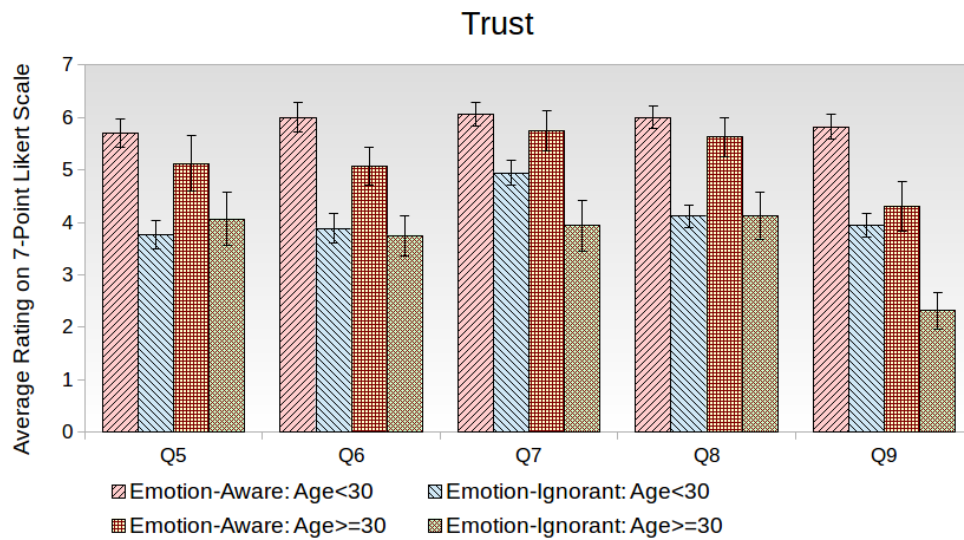


Figure 5.25: Impact of age on results of Likert scale questions related to trust.

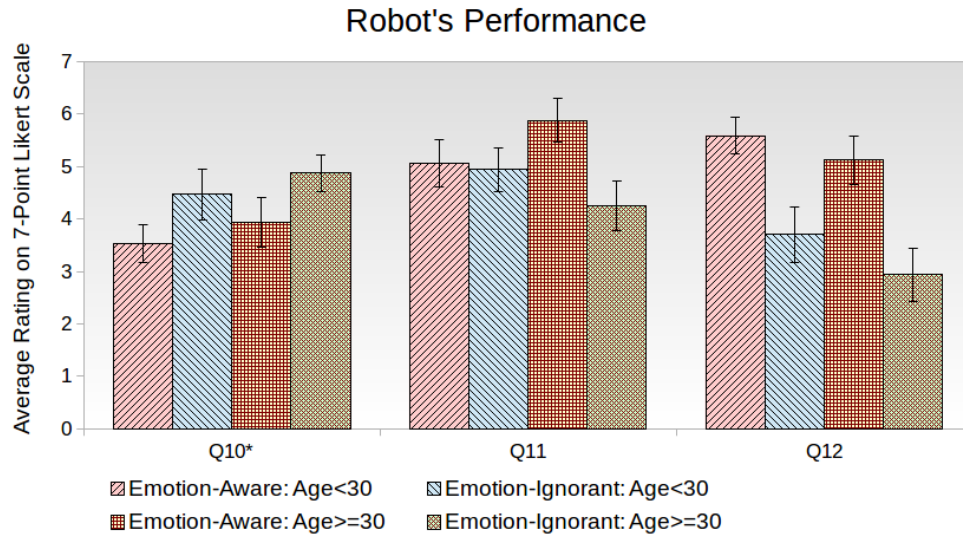


Figure 5.26: Impact of age on results of Likert scale questions related to performance.

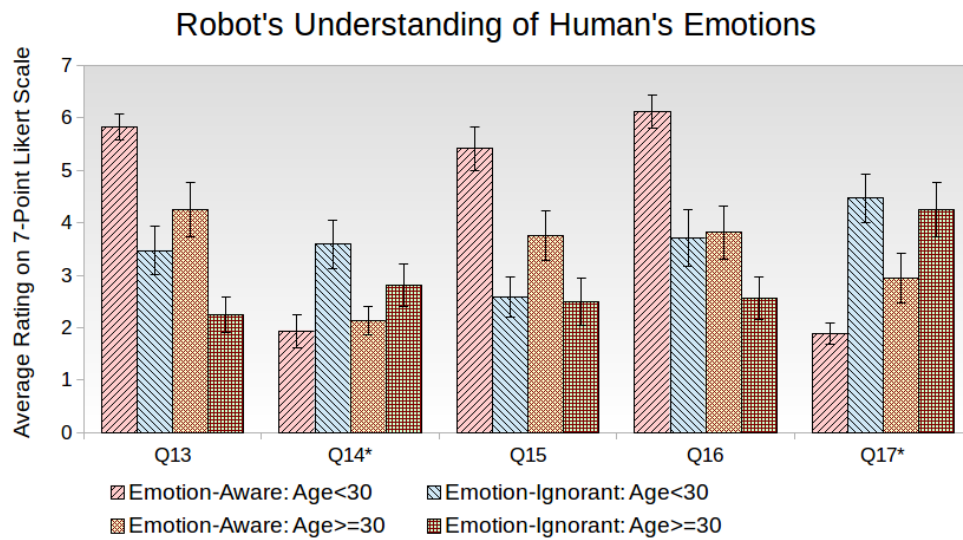


Figure 5.27: Impact of age on results of Likert scale questions related to robot's understanding of human's emotions.

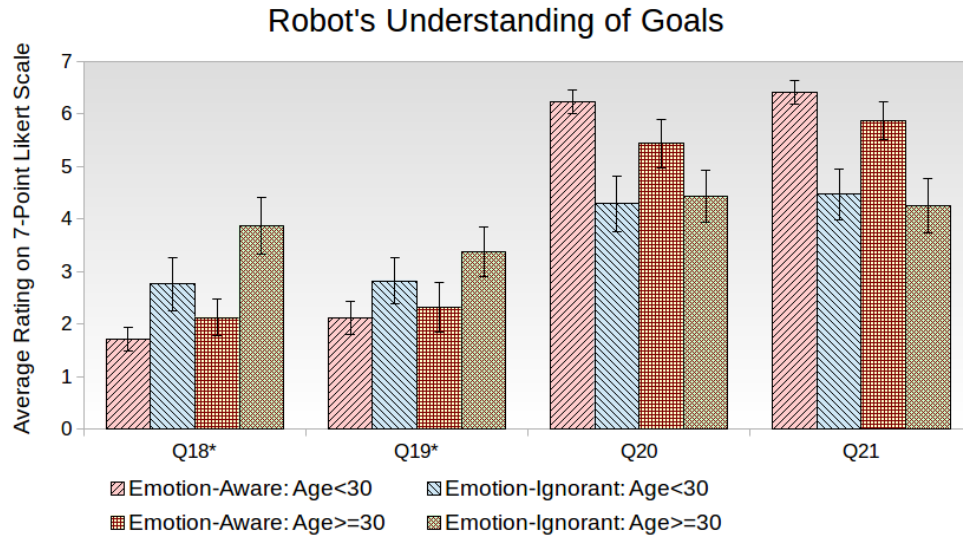


Figure 5.28: Impact of age on results of Likert scale questions related to robot's understanding of goals.

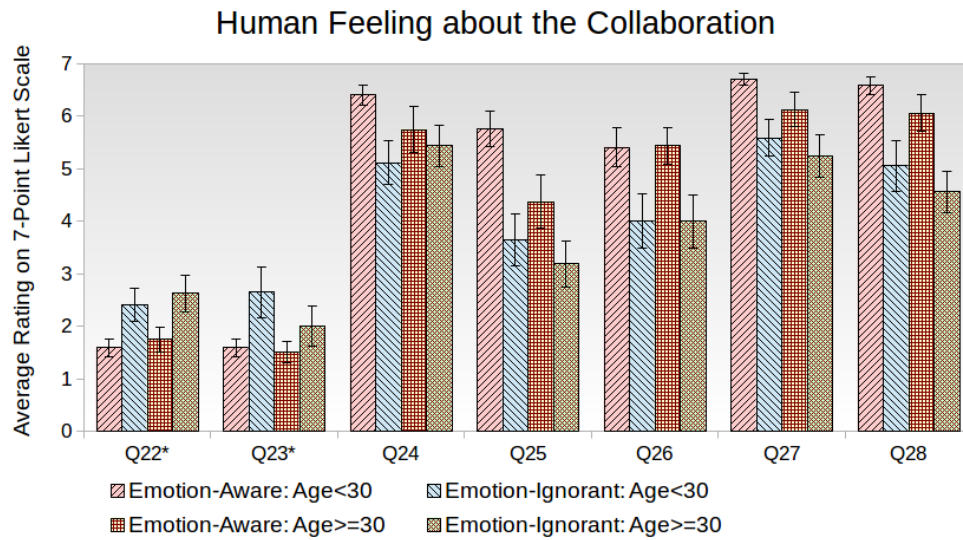


Figure 5.29: Impact of age on results of Likert scale questions related to human's feeling about collaboration.

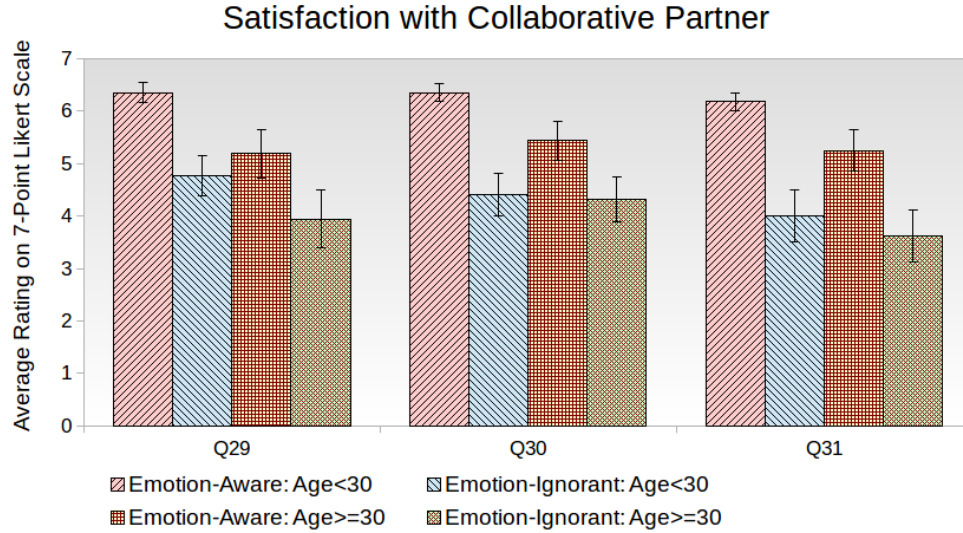


Figure 5.30: Impact of age on results of Likert scale questions related to satisfaction with collaborative partner.

5.2.5 Discussion

Based on the results, all participants prefer to work with the emotion-aware robot. Humans find the emotion-aware robot more likable and more trustworthy, as indicated in the Likert-scale responses and the open-ended questionnaire responses. Based on the responses, the emotional interaction with the robot can help create a sense of closeness and enjoyment that makes humans want to continue working with the robot.

The results also indicate that the emotion aware robot can better maintain a collaborative relationship. Both Likert-scale responses and Open-Ended Questionnaire responses indicate this. Humans felt a stronger sense of the robot's commitment to the collaboration, and greater understanding of their goals and emotions from the robot. Several open-ended responses also indicated that the robot was able to successfully motivate people and maintain their commitment to the collaboration, especially when tasks failed. Additionally, as shown in Section 5.2.4, humans rated the emotion-aware case much higher than the emotion-ignorant case when asked

which robot's decisions improved their performance, in essence acknowledging that their collaborator's (i.e. the robot's) decisions had a significant impact on their performance. As some of the open-ended responses indicated, successfully managing emotions within the collaboration can help keep the collaboration on track, and prevent distractions due to guilt and other negative emotions.

Finally, the emotion-aware robot developed a stronger sense of partnership through greater communication. The participants felt better understood by the emotion-aware robot, and felt that the goals were more mutually agreed-upon, refer to Section 5.2.4. As evidenced in the following response, the emotion-aware robot was successfully able to create a sense of partnership through its more open communication style: "Communication is very important. In the first run (i.e. emotion-aware) the robot states what tasks he is working on, it is clear and straight-forward. Also during the first run the robot cares about the human(me)'s feelings and cheers me up when I failed at the tasks, I think that could also improve efficiency of collaboration, because it would be more like a team or partnership."