Investigation of Robots' Perception Based on Agreement with Human Opinions

Kacper Nizielski, Alejandro Lopez, Emmanouil Zagoritis

LIACS, Leiden University
The Netherlands

nizielskik@gmail.com | alej.lopez.ruiz@gmail.com | zagoritisemmanuel@gmail.com

Abstract. The perception of the conversational partner, shaped by their agreement or disagreement and expressed through both verbal and non-verbal interactions, forms an impression of us that stays throughout the conversation. Our study aims to examine how robots are perceived based on varying levels of agreement with human opinions. To achieve this, three prototypes (agreeing, neutral, and disagreeing) will be programmed and deployed to have a conversation in the real world. Each robot will be supported with verbal cues (e.g. saying "yes") and non-verbal gestures (e.g. nodding).

1 Introduction

In the last year, artificial intelligence has advanced incredibly fast, enabling the field of robotics to design a vast amount of prototypes with an endless range of applications. Conversational robots, being one of the latter mentioned, are improving daily, and user experience is the core of it. With the significant improvements conversational AI has achieved, they can understand, process, and respond 4to Natural Language, the focus should be switched to the field of Human-Agent Interaction. The extent of how pleasant, intellectually, or bothering a conversation can be depends on the respective user the agent is interacting with and the approach followed, rather than the standard understanding of Natural Language the modern AI has.

Therefore, this paper will investigate how different conversational approaches towards human opinions influence the agent's perception. Agreement and disagreement can strongly shape a conversation, influencing autonomy, kindness, perceived trust, and much more. The variable of this study will be the agent's agreement with the opinions, varying from agreeing on everything, impartial to reflexive, and disagreeing. Thus, the paper aims to investigate how the latter variables impact the user's perception of the agent, delving into friendliness and intelligence.

2 Motivation and Related Work

Previous studies indicate that agreement and disagreement have an impact on a person's likability. Various conclusions have emerged across these studies: a person who disagrees is perceived as less competent or more biased, while others can modify their expressed attitudes when exposed to differing viewpoints [1]. It was also shown that individuals heavily rely on both non-verbal and verbal cues, which shape their views and assessments of others' opinions and trustworthiness [2].

Besides the studies on humans, research has also explored the dynamics between individuals and various agents, including robots. The experiments revealed for example that robots that expressed agreement - through nods and verbal confirmation - enhance people's likability and are perceived as friendlier during conversations [3-5]. Moreover, non-verbal cues significantly change a person's perception of the robot. Researchers found out that using non-verbal gestures supports and improves social engagement and helps robots be more “alive”, consequently boosting human satisfaction and likability [6].

Additionally, research into the autonomy and intelligence of robots shows that the robot's response to the user's opinion, either agreeing, neutral, or disagreeing, can also influence perceptions of its autonomy and intelligence. For example, a robot with neutral responses, agreeing with some opinions and disagreeing with others, can be perceived as more intelligent and autonomous [7].

3 Research Questions

The way social agents react or respond to users significantly impacts how their social qualities are perceived. Thanks to anthropomorphization, people generate a set of characteristics and abilities that an agent does not actually have. Abilities such as friendliness, intelligence, autonomy, and much more, essentially human traits, can be attributed to a robot based on its behavior. This can be exploited when designing social agents, taking advantage of the possibility of using human psychology to give a desired impression of oneself, the agent in this specific case. Therefore, this paper will evaluate how different degrees of agreement on human opinions shape how friendly, intelligent, and autonomous it appears to be, based on the user's perception.

Research Question:

Does a conversational robot that agrees with a user's opinions appear friendlier and more relatable than one that is neutral or disagrees, and how do these levels of agreement influence perceptions of the robot's intelligence, and autonomy?

This question will be addressed by programming three robot interaction styles (agreement, neutrality, and disagreement) each using verbal cues and supportive gestures to reinforce the interaction style.

3.1 Hypotheses/More Detailed Measurable Question

To be able to collect the data for answering the research question, the following hypothesis will be defined, which will be tested with the three robot models through a pilot study.

1. Friendliness Hypothesis

- A social agent that always agrees with the user's opinions will be perceived as friendlier and more approachable while a robot that disagrees will be sensed as mean and irritating.

2. Intelligence and Autonomy Hypothesis

- A social agent that carefully agrees with some opinions and disagrees with others will be perceived as more intelligent and autonomous than the ones that only agree or disagree.

These questions will be tested and further analyzed through the rest of the paper. We expect the users to feel a connection with a robot that is supposed to share their beliefs, and perceive a high level of autonomy when the robot appears to have different ones, therefore, its own beliefs.

4 Method

The study will be divided into two subsections. Firstly, we will do a feasibility study with a few participants to test our prototypes. Afterward, we will proceed with a pilot study using different versions of the agent and conduct it in a lab setting with a larger amount of participants. After the experiment, we are going to offer all the participants a Godspeed questionnaire to analyze their perception of the robot's intelligence. Additionally, we will interview the participants to gather qualitative data that will give an overall insight into the project's conclusion and will help us understand and analyze better the quantitative data when the final experiment is conducted.

4.1 Materials

In this study, we will be using the NAO robot to check our hypothesis and conduct our experiment. In order to represent the non-verbal cues, we will use various types of gestures that show how the robot agrees or disagrees, which the robot does through its actuators (e.g. nodding, thumbs up, and thumbs down). Moreover, there are going to be some prefabricated conversations and some random agreeing cues.

4.2 Experimental Setup/Approach

The first phase will be the feasibility study which will assess how varying levels of agreement affect perceptions of friendliness and intelligence in robot-human interactions. Three robot prototypes (agreeing, neutral, disagreeing) will be programmed with verbal and non-verbal cues. For this phase, six participants will have a single interaction with the robot in a lab setting to gauge initial responses and identify areas that need improvement before the main pilot study.

The second phase will be a pilot study with 12 participants, divided into three groups:

1. Agreeing Group: The robot agrees with the user.
2. Neutral Group: The robot remains objective.
3. Disagreeing Group: The robot disagrees with the user.

Before the interaction, participants will receive a brief description of the robot and one of the researchers will showcase a test conversation with the robot to demonstrate how the interaction will proceed. Later, participants will have a one-to-one conversation for approximately 7 minutes where the robot will introduce three discussion topics. For each discussion topic, there are three related subtopic questions, where the robot will ask the user's opinion and respond appropriately depending on its mode (agreeing, neutral, or disagreeing).

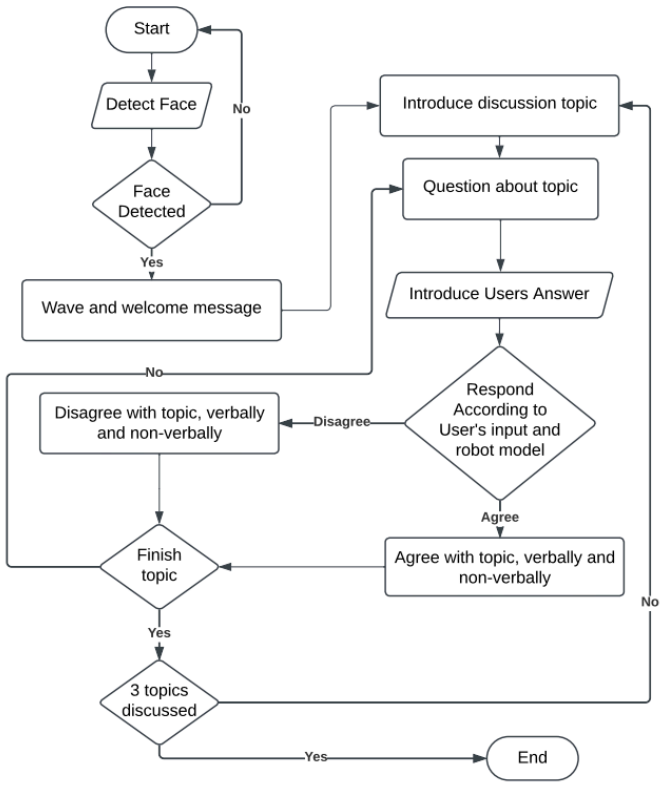
In the feasibility study, the autonomous approach will be tested, where the robot responds after the user completes their answer. If this approach proves ineffective, the Wizard of Oz approach will be adopted in the second phase. In this approach, the human operator will remotely control the robot, signaling it to respond at appropriate moments based on the user's input. This ensures consistency and accuracy of the real conversation, preventing silences or constant interruptions by the robot. This allows us to focus on participant responses without being constrained by technical limitations in the robot's autonomy.

The interaction between participants and the robot will be recorded so that it can be further analyzed in case anything goes unnoticed during the conversation. Afterward, participants will complete a Godspeed questionnaire, followed by an interview to gather the missing data and identify uncontrollable variables that might affect our results. This interview will also be recorded to capture new insights that were not observed during the process.

4.3 Interaction Diagram

Detect face & face detected – The robot will actively search for the user's face, entering this state until a face is detected. When detected, the robot will fix its gaze on the user. In terms of affordance, this indicates that the robot is conscious of the user's presence and ready to interact.

Wave and welcome message – Once the user is detected, the robot will wave and deliver a welcoming message, initiating a friendly and engaging interaction. This shows acknowledgment of the user's presence, reinforcing the idea that the robot is capable of understanding and responding to user feedback.



Introduce discussion topic – The robot will maintain three conversations with each user, chosen randomly from a pool. The robot will introduce and explain a discussion topic, giving the impression of being autonomous.

Question about topic – After introducing the topic, the robot will question the user, encouraging them to express their opinion. This step emphasizes the robot's ability to listen and interact meaningfully, fostering a sense of intelligence and engagement.

Introduce user answer – At this stage, the robot will process the user's response to determine their stance on the topic and prepare an appropriate reaction based on its programmed behavior.

Respond according to the user's input and robot model – Depending on the robot's programmed model, it will either agree or disagree with the user's opinion or maintain a neutral stance while presenting its perspective (which means that the answer will be the same regardless the user saying yes or no). While the agreement model would go to agree when the user says yes and disagree when the user says no, the disagreement model would do completely the opposite. This step is recursive and gets called for the first question (after introducing the topic) and for it three subtopics.

Agree/disagree with the topic, verbally and non-verbally – Based on the previous step, the robot will agree or disagree using gestures such as nodding or shaking its head, combined with verbal feedback. Positive cues (e.g., clapping) and invalidating ones (e.g., dismissive gestures) enhance the interaction's emotional depth, reinforcing the robot's stance.

Finish topic – The robot checks whether the predefined number of questions for the current topic has been discussed. If not, it continues; otherwise, it proceeds to the next topic. This ensures structured conversations without abrupt endings.

3 topics discussed – After discussing three topics, the interaction ends. If fewer topics are covered, the robot selects a new one, restarting the loop. This provides a clear conclusion, signaling the interaction's end while maintaining consistency.

4.4 Measures

Given that the study contains both feasibility study and pilot study methods, collecting qualitative data is appropriate. The variables that we need to measure are the likability and intelligence of an agent which we will achieve from a Godspeed questionnaire, which is a quantitive measure as there is a scale for the participant to choose. Additionally, there is going to be an interview from which we will get qualitative data by observing the answers of the participants.

5 Experimental results

The results of the study are presented below, divided by the three interaction models: Agreement, Disagreement, and Neutral. Each model's performance is detailed based on participant feedback and observed interactions during the study pilot.

1. **Agreement Model**

   All participants found the robot engaging, with particular praise for its expressive and immersive gestures, such as changing eye color and nodding. Participants frequently remarked that the robot's behavior felt natural and was aligned with their expectations, with some describing the robot as almost human-like in its interactions. Some noted that some occasional repetitiveness occurred in responses during conversation. Also, some of the non-verbal cues were exaggerated, like clapping.

   The level of connection with the robot varied among participants. While most of them felt some sense of connection, due to their engaging behavior, there were some participants who had not a feeling of any bond. Participants suggested areas for improvement, such as more enhanced expressiveness from the robot's side introducing more dynamic non-verbal cues (eye movement), and improving the sensitivity of auditory sensors to better capture input from the user.

   The Agreement Model strongly supported the Friendliness Hypothesis, as it was the most positively received and described as friendly and approachable. However, it provided only limited support for the Intelligence and Autonomy Hypothesis. While participants appreciated the robot's engaging behavior, they did not emphasize any signs of intelligence or autonomy.
   Overall, the Agreement Model was highly appreciated for its positive reinforcement and natural use of gestures, which were pleasant for participants. Although the model had many positive aspects, its weaknesses included occasional repetitiveness and a lack of dynamic expressiveness in some cases, suggesting room for improvement in the future.

2. **Disagreement Model**

   Some participants noted that the robot's disagreement behavior encouraged them to think deeper and actively participate by challenging their opinions. They also felt that the conversation was more dynamic and engaging due to contradictions, which made participants more interested in explaining their points more thoroughly. However, despite these positive aspects, due to the lack of common ground and frequent disagreements, all participants found the robot's disagreement behavior frustrating, which ruined the overall conversation. Participants did not report feeling a strong connection with the robot. Some participants mentioned that they felt only a superficial connection, by feeling frustration with a robot's stance. One participant likened the disagreements to a personality trait, appreciating the perception of independence, but even in this case, there was no meaningful bond established.

The disagreement Model provided limited support for the Friendliness Hypothesis. While participants said that interaction was engaging, they did not perceive the robot as friendly or approachable. In contrast, the model offered moderate support for the Intelligence and Autonomy Hypothesis. The robot's disagreement behavior contributed to a perception of independence and autonomy, but this was perceived as artificial rather than intelligent.

Overall, the Disagreement Model encouraged active participation and critical thinking by introducing contrasting opinions, creating a sense of autonomy. However, its persistent disagreement felt unnatural, emotionally cold, and repetitive, reducing likability. Participants suggested making the disagreements more polite and balanced, as well as incorporating occasional agreements to increase naturalness and relatability.

3. Neutral Model

Participants found the robot's responses logical and structured, which led some to perceive the robot as intelligent. They also considered the robot engaging but noted that its responses sometimes felt abrupt, and there was minor repetitiveness during the interaction.
All participants reported a limited emotional connection with the robot, attributing this to its neutral tone and lack of expressiveness. The robot's neutral stance made it harder for participants to form a deeper bond, as the robot's interactions felt more mechanical than emotionally engaging. Even when the robot's responses were rational, they lacked the warmth often associated with human-like interactions.

The Neutral Model provided limited support for the Friendliness Hypothesis. While the robot was seen as competent, participants felt the interaction lacked emotional warmth, making it less approachable and friendly. The lack of empathetic responses further strengthens this perception, making the robot more distant. In terms of the Intelligence and Autonomy Hypothesis, the model provided moderate support. The robot's structured responses were regarded as intelligent, but its lack of emotional engagement or variation in responses led to a perception of reduced autonomy, as it seemed to operate more like a programmed system than an agent with independent thought.

Overall, while some participants found the robot's responses to be intelligent, the emotional flatness and mechanical tone of the interaction reduced its perceived friendliness. The lack of dynamic engagement and expressiveness left the robot feeling distant, preventing a more emotionally connected experience. The overall interaction felt more like an exchange of information rather than a meaningful, human-like conversation.

6 Discussion

The results of this study gave insights into how Agreement, Disagreement, and Neutral interaction models shape perceptions of a robot's likability, friendliness, intelligence, and autonomy. These results underscore how different interaction styles influence user perceptions and extend the insights found in prior research.

Previous studies emphasize that agreement and disagreement influence perceptions of likability and competence [1, 3]. The Agreement Model aligns with these findings by showing that gestures such as nodding or showing a thumbs-up can enhance friendliness and approachability [4, 5]. Participants, as mentioned, preferred these non-verbal behaviors but mentioned the exaggerated or overly frequent gestures as cues that spoiled the perception of naturalness and authenticity [6], emphasizing the need for dynamic non-verbal interactions that feel spontaneous rather than preprogrammed.

The Disagreement Model encouraged the participants to justify their beliefs and engage in deeper thoughts aligned with prior research suggesting that disagreement can enhance cognitive engagement [1]. However, persistent disagreement with the user's responses frustrated participants, which is a challenge identified in a prior study [2], created because of the robot's lack of balance or politeness. Furthermore, the repetitive nature of the robot's disagreement responses reduced its perceived intelligence, as it gave the impression of being firmly programmed rather than adaptively responsive. Introducing more varied and contextually appropriate disagreements could enhance the robot's autonomy without sacrificing likability [7].

The Neutral Model results showed responses that were logical and structured. Participants recognized these responses as intelligent, just like findings in earlier studies about the role of logical consequence in shaping perceptions [5]. Despite that, the robot with this model had the absence of warmth and expressiveness, which led to participants being disconnected and perceiving the robot as mechanical and separate. This aligns with earlier research suggesting that warmth and non-verbal cues are critical for establishing a meaningful connection with conversational agents [6]. Although neutrality may create a perception of the robot being competent, its emotional flatness limits the potential for deeper relational engagement. Connecting more empathetic behaviors to the robot can create a more powerful sense of engagement during the conversation. This result underscores the importance of connecting empathetic cues with neutral communication styles to create a balance between logic and emotional resonance.

Through all the models, the most important improvement is creating more natural varied interactions. Participants repeatedly emphasized that overly stiff cues or repetitive responses spoiled their perception of the robot's autonomy and intelligence. This aligns with studies highlighting the significance of adaptability in conversational agents for creating more engaging and lifelike interactions [7]. Furthermore, more dynamic non-verbal behaviors, such as varied gestures and facial expressions, were shown to play a vital role in supporting engagement and likability.

7 Conclusion and further research

In conclusion, the results highlight the connection between friendliness, autonomy, and intelligence. The Agreement Model created a sense of friendliness, but it fell short of projecting autonomy. Conversely, The Disagreement Model succeeded in suggesting independence but at the cost of likability. The Neutral Model was in a middle ground but lacked the emotional connection. These insights suggest that the best interaction style would combine elements of all three models balancing through agreement, disagreement, and neutrality.

There is still significant work to be done in determining which model performs best or if the connection of the three models is the best one. As previously mentioned, non-verbal cues play a vital role in interactions, but these were quite limited with the NAO robot. The robot was capable of performing only basic gestures, such as clapping or raising hands. Participants, however, suggested that eye movement would enhance its expressiveness. Unfortunately, this is not possible with the NAO robot, as it only features two lenses surrounded by LED lights.z

Furthermore, there is a lot of room for improvement to exploit this experiment's potential. With the proper equipment and preparation, this experiment should be conducted outside a laboratory environment, allowing it to feel as natural as possible The robot should be able to hold a proper conversation with any random participant and have a set of non-verbal cues that increase engagement more than the actuals, such as the ones mentioned in the paragraph above. Also, ideally, the users should be the ones able to propose the conversation topics, making the robot seem more intelligent as it can adapt to any topic. This could be accomplished with a better robot model than the NAO, working with a large language model.

References

[1] J. Minson, “The psychology of disagreement with Julia Minson”, Harvard Kennedy School, 2021. [Online] Available: https://www.hks.harvard.edu/wiener-conference-calls/julia-minson

[2] J. A. Hall, M. S. Mast, and T. V. West. “Accurate interpersonal perception: Many traditions, one topic”. The Social Psychology of Perceiving Others Accurately, Cambridge University Press. 2016. pp. 3–22. doi: 10.1017/CBO9781316181959.001.

[3] H. Admoni. “Nonverbal Communication for Human-Robot Interaction”. ACM Transactions on Embedded Computing Systems. 2014. pp. 1-3. doi: 10.1145/0000000.0000000

[4] H. Admoni. “Nonverbal communication in socially assistive human-robot interaction”. AI Matters. 2016. pp. 4-9. doi: 10.1145/3008665.3008669.

[5] Yale Social Robotics Lab. “Nonverbal communication in human-robot interaction: Henny Admoni”. Yale University. [Online] Available: https://scazlab.yale.edu/nonverbal-communication-human-robot-interaction-henny-admoni

[6] J. Urakami and K. Seaborn. “Nonverbal cues in human-robot interaction: A communication studies perspective”. 2023. doi: 10.48550/arXiv.2304.11293

[7] T. Kanda and H. Ishiguro. “Human-Robot Interaction in Social Robots”. Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication. 2013. pp. 698-703. doi: 0.1201/b130041