

# What Am I? Evaluating the Effect of Language Fluency and Task Competency on the Perception of a Social Robot

Shahira Ali, Haley N. Green, and Tariq Iqbal

**Abstract**—Recent advancements in robot capabilities have enabled them to interact with people in various human-social environments (HSEs). In many of these environments, the perception of the robot often depends on its capabilities, e.g., task competency, language fluency, etc. To enable fluent human-robot interaction (HRI) in HSEs, it is crucial to understand the impact of these capabilities on the perception of the robot. Although many works have investigated the effects of various robot capabilities on the robot’s perception separately, in this paper, we present a large-scale HRI study ( $n = 60$ ) to investigate the combined impact of both language fluency and task competency on the perception of a robot. The results suggest that while language fluency may play a more significant role than task competency in the perception of the verbal competency of a robot, both language fluency and task competency contribute to the perception of the intelligence and reliability of the robot. The results also indicate that task competency may play a more significant role than language fluency in the perception of meeting expectations and being a good teammate. The findings of this study highlight the relationship between language fluency and task competency in the context of social HRI and will enable the development of more intelligent robots in the future.

## I. INTRODUCTION

The integration of robots to interact and work alongside people has become a growing area of research [1]–[7]. Advancements in robotics over the last few decades have enabled them to perform complex tasks and interact with people across various human-social environments (HSEs) [8]–[13]. As such, there has been an increase in research understanding how robots are perceived in various scenarios based on their roles [14]–[20].

One particular capability of robots that has widely been explored in the context of social human-robot interaction (HRI) is their verbal communication skills [21]–[26]. Along this research direction, many works have focused on investigating the communication abilities of robots, and their effectiveness in various task settings [7], [17], [22], [27]. Another work evaluated interpersonal communication and relationship building between a robot and a human partner and found that verbal communication capabilities had a positive impact on the social perception of the robot as well as can strengthen the bond between the collaborative pair [27]. Similarly, another study found that verbal communication positively impacts perceptions of friendliness and social presence in robots [28]. Given the critical role verbal communication plays in HRI, many works in HRI have focused on making communication as natural and fluent as possible [22], [23].

The authors are with the School of Engineering and Applied Science, University of Virginia, Charlottesville, VA, USA. [{sma6zw,hng9vf,tiqbal}@virginia.edu](mailto:{sma6zw,hng9vf,tiqbal}@virginia.edu).



Fig. 1. A NAO robot plays the “What Am I?” game with a human. The participant sits opposite the robot during the game and guesses the identity of the animal that the robot is assuming from the list provided to them.

In human-human interactions, among various aspects of verbal communication, language fluency particularly plays a significant role in building a perception of a person. For example, some studies indicate that participants hold negative biases and stereotypes against non-native speakers [29]–[31]. Although the impact of language fluency on the perception of a person in human-human interaction has been considerably studied, there has not been any work exploring the impact of language fluency on the perception of the robot. As robots are now expected to interact with people in various social scenarios, it’s crucial to explore how verbal communication, specifically language fluency, may play a role in HRI.

Along with verbal communication being a significant part of interactions, robots are also expected to interact with people and perform various tasks with them. In these scenarios, the task competency of the robot also affects how people perceive the robot. Several studies have investigated varying task competency and its influence on the perception of robots [32]–[38]. For example, robots that were interruptive during tasks were perceived as being less task competent than the robots that completed the tasks without interleaving [35], [38]. Additionally, results from some works indicate how human agents perceived an erroneous robot as less trustworthy and less reliable than a competent robot in a collaborative setting [39].

Although many works focus on investigating the impact of a robot’s verbal communication ability and task competency on the perception of the robot separately, to the best of our knowledge, no research to date has studied the combined impact of language fluency and task competency on the perception of a social robot. In this work, we are interested in exploring how varying language fluency and task competency in a social robot will affect people’s perceptions of the robot. Specifically, we aim to investigate

the impact of the robot's language fluency and task competency (i.e., fluent-competent, fluent-incompetent, disfluent-competent, and disfluent-incompetent conditions) on the perceptions of the robot's verbal competence, intelligence, and reliability along with whether the robot meeting expectations, being a good teammate, and if the participant is willing to work with the robot again. To investigate that, we have designed a human-robot interaction scenario modeled after the children's guessing game, "What Am I?", and conducted a large-scale HRI study ( $n = 60$ ). In this task, a NAO robot describes various characteristics of an animal, and the human participant is prompted to guess what animal the robot is referring to. We vary the language fluency of the robot by allowing it to speak at a native or a non-native level and the task competency by allowing the robot to correctly match or fail to identify the animal.

The results from the study suggest that while language fluency may play a greater role than task competency in the perception of verbal competence of a robot, both language fluency and task competency contribute to the perception of intelligence and reliability of the robot. The results also indicate that task competency may play a greater role than language fluency in the perception of meeting expectations and being a good teammate. These findings will allow us to have a deeper understanding of the relationship between language fluency and task competency on the perception of a robot and to develop ways to reduce non-native language biases and promote greater understanding and respect for linguistic diversity in the context of HRI.

## II. BACKGROUND

### A. Language Fluency

The definition of fluency in the context of spoken language in human-human interactions (HHI) has long been debated and studied. For example, Lennon [40] differentiates between two approaches of fluency. The first is a holistic, broad sense in which fluency refers to global oral proficiency. The second portrays fluency in a much narrower sense and relies on the efficient and effortless planning and production of speech, such as speaking speed or grammatical accuracy. Furthermore, language fluency has been linked to higher levels of perceived proficiency and is "best conceived of as fast, smooth, and accurate performance" [41], [42]. However, there is a greater risk of confounding measures when multiple measures are used to define and examine fluency [43]. Given the limited literature on definitive descriptions and classifications of language fluency in HHI, it is not surprising that language fluency in HRI has yet to be adequately explored. In this study, we will be referring to language fluency in Lennon's narrow sense and focus only on one measure of oral proficiency: *grammatical accuracy*.

### B. Language Bias

The study of language bias and perception has become increasingly important as globalization has led to a more diverse and multicultural society. One area of interest is the perception of non-native language speakers by native

language speakers. For example, several studies have been conducted to investigate the extent of bias, and the factors that influence perceptions of non-native speakers [29], [30], [44]–[48]. Other studies depict how students express a preference for native English-speaking teachers and negatively judge the teaching skills of non-native English teachers [31], [49], [50]. Another study found non-native speakers to have lower perceived intelligence and trustworthiness by native speakers [51]. There is adequate literature surrounding native speaker bias and perceptions of non-native speakers in human-human interactions. However, little research has been done surrounding the impact of a person's linguistic abilities on their perceptions of a robot. A native speaker is someone who speaks that language with native fluency [52]. To reduce the bias of participants' linguistic background, in this work, we required all participants to be monolingual in English.

### C. Language in Robots

Verbal communication has been a crucial area of research in robotics, as it enables robots to interact with humans in a more natural and intuitive way [21]–[23], [25]. Language plays an important role in human-robot interaction, as it allows robots to understand and respond to human commands, convey information, and establish social relationships with humans [53]–[55]. The focus of several works has been to improve perceptions of trust in robots by employing deep learning speech and automatically generated explanations [56], [57]. With the importance of verbal communication and natural language established in HRI, it is important to understand how varying levels of language fluencies can impact perceptions of that robot.

### D. Task Competency

Previous research in HRI has explored various aspects of robot behavior, including task performance, and their impact on the perception of the robot [24], [26]. For example, Salem et al. found that participants rated a robot as more reliable and trustworthy when it performed a task competently compared to when it made errors, suggesting that task competency influences perceptions in robots [39]. In a similar vein, Carter et al. found that participants rated robots who did not interrupt during a task as more competent than robots that did interrupt the collaborative task [38]. Furthermore, Clair et al. employed verbal feedback in a human-robot collaboration task and found that ratings of team performance and the robot as a teammate were improved [58]. Although some work explored the interaction between verbal messages and task competency, there is a gap in the exploration of the effect of language fluency and task competency in HRI. Therefore, in this paper, we aim to examine the effects of varying task competency paired with varying language fluency on the perceptions of a social robot partner.

## III. RESEARCH QUESTIONS

To examine the impact of varying language fluency and task competency on the perception of a social robot, we have designed four experimental conditions. In particular, we aim

to investigate whether perceptions vary for a robot that is: (1) fluent and task competent; (2) fluent and task incompetent; (3) disfluent and task competent; or (4) disfluent and task incompetent. We refer to these four conditions as the varying robot conditions. Building on the literature, we want to investigate the impact of these varying robot conditions on the general perception of the robot (verbal competency, intelligence, reliability), the expectations being met, the robot as a teammate, and participants' willingness to work with the robot again.

Therefore, we aimed to address the following four research questions:

- **RQ1:** How do the perceptions (verbal competency, intelligence, and reliability) of the robot partner differ across the varying robot conditions?
- **RQ2:** How do the perceptions of the participants' expectations being met differ across the varying robot conditions?
- **RQ3:** How do the perceptions of the robot as a teammate differ across the varying robot conditions?
- **RQ4:** How does the willingness to work with the robot differ across the varying robot conditions?

#### IV. METHODS

##### A. What Am I? Study Design

To address these research questions, we designed a human-robot interactive scenario loosely based on the guessing game, "What Am I?". This game consists of the interaction partner prompting the robot to guess the identity of an animal that the robot is assuming after hearing a few characteristics of that animal from the robot. The robot was positioned on a small table to face the human agent during the entire interaction. The participant was provided with a list of animals and their associated characteristics, which they could reference throughout the game. The participant was either assigned to a fluent robot, which spoke English at a native fluency level, or a disfluent robot, which spoke English at a non-native fluency level, for the entire duration of the game. We define fluency in detail in the following section.

To initiate the game, the robot greets the player and asks if they are ready to begin the round. After the partner affirms, the robot states the characteristics of the animal whose identity it assumed and asks the participant, "What am I?" (fluent robot) or "What I am?" (disfluent robot). Then, the participant identifies the animal, and the robot either correctly confirms the accuracy of the identity or incorrectly rejects the participant's selection and states the wrong animal for the characteristics it assumed. The robot then continues to the next round and repeats this process with a different animal identity until a total of three rounds are completed. At the conclusion of the third round, the robot states the game has ended and thanks the human partner for playing. Table I shows a sample script of the varying robot interactions.

##### B. Language Fluency and Task Competency Classification

**Language Fluency:** There are many factors that can impact the spoken language fluency of a non-native speaker and the

TABLE I  
SAMPLE SCRIPT FOR THE FOUR VARYING ROBOT CONDITIONS.

	Fluent	Disfluent
Task Competent	NAO: I like bananas and I climb trees. What am I? Human: Monkey NAO: Yes, that is correct. Let's go to the next round.	NAO: Me bananas and climb tree. What I am? Human: Monkey NAO: Yes, correct. We go next round.
Task Incompetent	NAO: I like bananas and I climb trees. What am I? Human: Monkey NAO: No, I am a rabbit. Let's go to the next round.	NAO: Me bananas and climb tree. What I am? Human: Monkey NAO: No, me rabbit. We go next round.

perception of their language fluency such as stutter, accent, hesitation, intonation, phonological processes, pronunciation rules, and ungrammatical words or sentence structures, etc. [40], [42], [59], [60]. For this study, we incorporated grammatical accuracy in a speech to represent the fluent robot partner. In order to accurately depict disfluency in the robot, we need to define speech disfluency. A possible taxonomy to understand speech disfluencies is by distinguishing speech disfluencies into two groups: disfluencies that stem from uncertainty and errors or error-type disfluencies (ETDs) [61]. This taxonomy lists the principle measures of uncertainty-related speech disfluencies to include factors such as hesitations and repetition while error-type disfluencies encompass measures such as grammatical errors and contamination. In this work, we interpreted the root of speech disfluency to be ETDs. Specifically, we incorporated grammatical errors in a speech to represent the disfluent robot partner. We also chose this factor because grammatical errors can be made easily decipherable to participants in robot speech, and we wanted to avoid disfluency measures that could be interpreted as inherent to the robot. For example, stutter or hesitation could have been incorrectly attributed to or perceived as malfunctions or natural behavior of the robot. The types of grammatical errors implemented in the disfluent robot's dialogue included inaccurate subject-verb agreement, incorrect singular/plural agreement, wrong pronoun usage (subjective vs objective), incorrect word form, and lack of article and preposition usage. These errors were based on the most common grammatical errors made by non-native English speakers [62]–[64]. A sample script of the disfluent robot interaction is included in Table I.

**Task competency:** Along with verbal language fluency, this study also looked at the perceptions of two types of task competencies by the robot: task competent (accurate task completion) and task incompetent (inaccurate task completion). The *task competent condition* was demonstrated in the game by the robot correctly affirming the animal that the human player guessed to be the robot's assumed identity. The *task incompetent condition* was represented in the game by the robot incorrectly rejecting the animal that the human player guessed to be the robot's assumed identity and stating a different incorrect animal as the answer. A sample script of the task competent and task incompetent robot interactions are included in Table I.

### C. Participant Classification

To reduce the participants' linguistic bias in the perception of the language fluency and task competency of the robot, we only chose to select monolingual participants to participate in this study. The monolingual categorization was based on participants' responses on a pre-task survey that collected biographical, demographic, and linguistic information and was completed upon arrival. The linguistic portion comprised two questions where the first asked how many languages the participant had verbal native fluency in, and the second followed up by asking to list the language(s) from the previous question. Native fluency was defined to the participants on the survey as being a native language that does not contain unnatural grammatical errors, pauses, stutters, repetitions, or self-corrections when speaking in the language. The participants of this study identified themselves as monolingual native English speakers. In order to be classified as a monolingual native speaker, the participant had to self-report their linguistic abilities and answer the first question with the number "1" and the second question with the language "English".

### D. The Robot

In our study, the SoftBank Robotics' NAO humanoid robot was used. To capture the robot's nonverbal gestures, we utilized the expressive behavior modules in SoftBank's Choregraphe Suite and for the robot's verbal performance, we worked with Amazon Polly's text-to-speech platform [65]. The decision to use Amazon Polly over Choegraphe's text-to-speech option was due to Amazon Polly's robust annunciation and timing capabilities. To mitigate potential bias with the perceptions of the robot's gender, we use the gender-neutral "ivy" voice. Then, 18 sound clips of the fluent, disfluent, accurate task completion, and inaccurate task completion dialogue were added to the Choregraphe program and matched with some basic animation behaviors. In Choregraphe, we also utilized the speech recognition modules so that the robot could react accordingly to the participant's responses.

## V. EXPERIMENT

### A. Procedure

Participants reviewed a study information document for consent along with the task instructions and were asked to complete a brief demographic survey. Each participant was randomly assigned to one of four robot conditions. Next, the participants played three rounds of the "What Am I?" game with the NAO robot. In each round, the robot would assume the identity of an animal and give a description of that animal to the participant, who would then have to guess the animal's identity based on the given description. After playing all three rounds, participants were asked to rate how much they agreed with the robot being verbally competent, intelligent, reliable, a good teammate, meeting their expectations, and willing to work with the robot again on a five-point Likert scale ranging from "strongly disagree" (1) to "strongly agree" (5). The entire study took

approximately 15 minutes. At the conclusion of the study, participants were debriefed and compensated.

### B. Participants

A total of 60 adults participated in the study (70.0% female ( $n = 42$ ), 30.0% male ( $n = 18$ )). The mean age of participants was 21 years ( $SD = 2.75$ ). All participants were required to be English speakers, available to participate in-person, at least 18 years of age or older, and monolingual. The participants consisted of undergraduate and graduate students, and high school and college graduates. During the study, participants also reported their experience with robots in general ( $M = 1.42$ ,  $SD = 0.740$ ). Participants were compensated with a \$5 gift card for participating in the study. The study was approved by the Institutional Review Board.

### C. Measures

After the game, participants were asked to rate perceptions of the robot's verbal competency, intelligence, and reliability. We also had participants reflect on whether the NAO robot was a good teammate, rate their willingness to work with the robot again, and express whether the robot met their expectations on the Likert scale described in Sect. V-A.

## VI. RESULTS AND DISCUSSION

To address our research questions, we evaluated the effects of varying language fluency and task competency on the perceptions of a robot partner. A series of multivariate ANOVAs were conducted with four groups that make up the varying robot conditions as independent variables and six dependent variables (i.e., perceptions of verbal competence, intelligence, reliability, expectations being met, the robot as a teammate, and willingness to work with the robot again). For the analysis, an inspection of the data and Levene's test provided no strong evidence against the assumption of constant variance. Furthermore, we used Tukey's Honestly Significant Difference (HSD) test for post-hoc comparisons. The multivariate analysis highlighted a statistically significant interaction effect between varying robot conditions on the combined perceptions of the robot partner, ( $F(18, 144) = 4.999$ ,  $p < .001$ , Wilks'  $\lambda = .255$ ,  $\eta^2 = .366$ ).

### RQ1: Effects of Varying Robot Conditions on Perceptions

1) *Verbal Competence:* Separate univariate ANOVAs on the dependent variables revealed significant effects of varying robot conditions on perceptions of verbal competence, ( $F(3, 56) = 16.46$ ,  $p < .001$ ,  $\eta^2 = .469$ ). The results (Fig. 2-A) suggest participants in the fluent-competent condition rated verbal competence ( $M = 4.53$ ,  $SD = 1.060$ ) higher than participants in the disfluent-competent group ( $M = 2.67$ ,  $SD = 1.397$ ), ( $p < .001$ ). Participants in the fluent-competent condition rated verbal competence ( $M = 4.53$ ,  $SD = 1.060$ ) higher than participants in the disfluent-incompetent group ( $M = 2.00$ ,  $SD = 1.604$ ), ( $p < .001$ ). Participants in the fluent-incompetent condition ( $M = 4.60$ ,  $SD = 0.828$ ) rated verbal competence higher than participants in the disfluent-competent group ( $M = 2.67$ ,  $SD = 1.397$ ), ( $p < .001$ ). Participants in the fluent-incompetent condition rated verbal

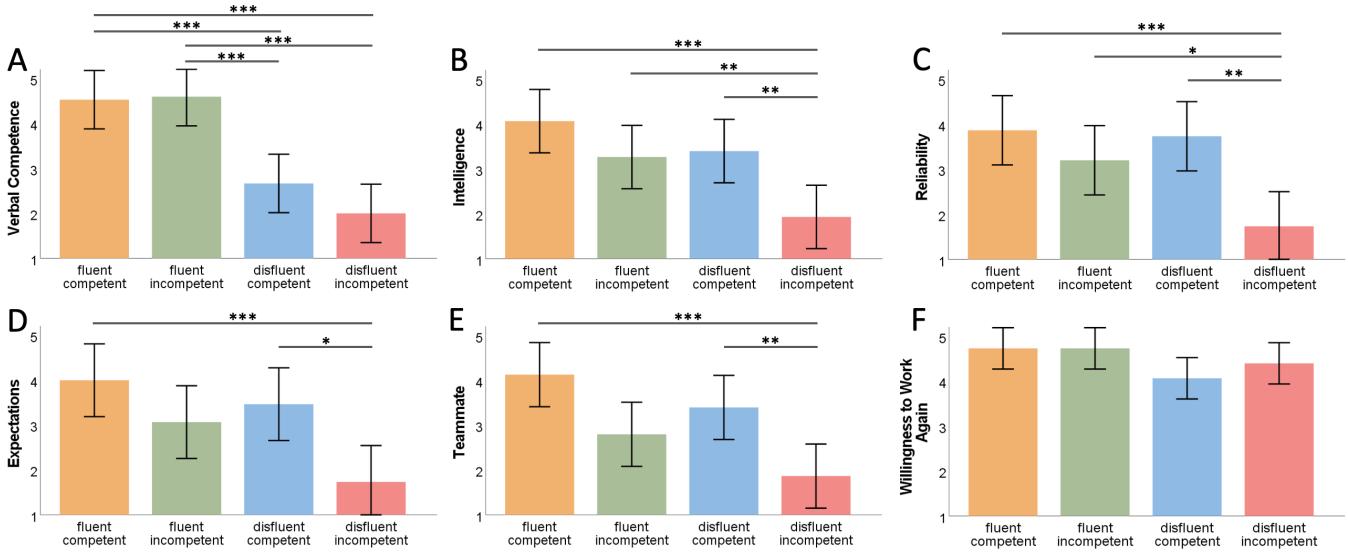


Fig. 2. Bar graphs showing participants' perceptions of verbal competence, intelligence, and reliability as well as their ratings of expectations being met, the robot as a teammate, and their willingness to work with the robot again across four varying robot conditions (the robot being fluent and task competent, fluent and task incompetent, disfluent and task competent, and disfluent and task incompetent). Error bars 95% CI. The significant values are shown in \* ( $* = p < .05$ , \*\*  $= p < .01$ , \*\*\*  $= p < .001$ ).

competence ( $M = 4.60$ ,  $SD = 0.828$ ) higher than participants in the disfluent-incompetent group ( $M = 2.00$ ,  $SD = 1.604$ ), ( $p < .001$ ).

**Discussion:** The results indicate that participants found both the fluent robots (fluent-competent and fluent-incompetent) to be significantly more verbally competent than the disfluent robots (disfluent-competent and disfluent-incompetent). This was expected as the disfluent robots made grammatical errors, which may influence the perception of the verbal competency of those robots. The results may indicate that language fluency plays a greater role than task competency in the perception of verbal competency of a robot.

2) *Intelligence:* Separate univariate ANOVAs on the dependent variables revealed significant effects of varying robot conditions on perceptions of intelligence ( $F(3,56) = 6.425$ ,  $p < .001$ ,  $\eta^2 = .256$ ). The results (Fig. 2-B) suggest participants in the fluent-competent condition rated intelligence ( $M = 4.07$ ,  $SD = 1.534$ ) higher than participants in the disfluent-incompetent condition ( $M = 1.93$ ,  $SD = 1.534$ ), ( $p < .001$ ). Additionally, participants in the fluent-incompetent condition rated intelligence ( $M = 3.27$ ,  $SD = 1.163$ ) higher than participants in the disfluent-incompetent condition ( $M = 1.93$ ,  $SD = 1.534$ ), ( $p = .047$ ). Participants in the disfluent-competent condition rated intelligence ( $M = 3.40$ ,  $SD = 1.183$ ) higher than participants in the disfluent-incompetent condition ( $M = 1.93$ ,  $SD = 1.534$ ), ( $p = .024$ ).

**Discussion:** The results suggest that the participants perceived the fluent-competent, fluent-incompetent, and disfluent-competent robots to be more intelligent than the disfluent-incompetent robot. It was expected for the fluent-competent robot to be perceived as more intelligent than the disfluent-incompetent robot due to neither language fluency or task competency being compromised in the former condition. However, in the fluent-incompetent and disfluent-competent conditions, at least one of the variables (language

fluency or task competency) was compromised, yet robots in both conditions were still perceived to be more intelligent than the disfluent-incompetent robot, which had both variables compromised. These results may suggest that both language fluency and task competency have an impact on the perceived intelligence of the robot.

3) *Reliability:* Separate univariate ANOVAs on the dependent variables revealed significant effects of varying robot conditions on perceptions of reliability ( $F(3,56) = 6.463$ ,  $p < .001$ ,  $\eta^2 = .257$ ). The results (Fig. 2-C) suggest participants in the fluent-competent condition rated reliability ( $M = 3.87$ ,  $SD = 1.807$ ) higher than participants in the disfluent-incompetent condition ( $M = 1.73$ ,  $SD = 1.534$ ), ( $p = .001$ ). Additionally, participants in the fluent-incompetent condition rated reliability ( $M = 3.20$ ,  $SD = 1.207$ ) higher than participants in the disfluent-incompetent group ( $M = 1.73$ ,  $SD = 1.534$ ), ( $p = .044$ ). Furthermore, participants in the disfluent-competent condition rated reliability ( $M = 3.73$ ,  $SD = 1.335$ ) higher than participants in the disfluent-incompetent condition ( $M = 1.73$ ,  $SD = 1.534$ ), ( $p = .003$ ).

**Discussion:** The results indicate that participants in the fluent-competent, fluent-incompetent, and disfluent-competent conditions perceived the robot to be significantly more reliable than the participants in the disfluent-incompetent condition. Similarly to the intelligence results, it was expected for the fluent-competent robot to be perceived as more reliable than the disfluent-incompetent robot due to neither language fluency or task competency being compromised in the former condition. However, in the fluent-incompetent and disfluent-competent conditions, at least one of the variables (language fluency or task competency) was compromised, yet robots in both conditions were still perceived to be more reliable than the disfluent-incompetent robot, which had both variables compromised. The results may imply that both language fluency and task competency

have an impact on the perceived reliability of the robot.

#### *RQ2: Effects of Varying Robot Conditions on Participants' Expectations Being Met*

Separate univariate ANOVAs on the dependent variables revealed significant differences between the varying robot conditions on ratings of expectations being met ( $F(3, 56) = 5.716, p = .002, \eta^2 = .234$ ). Specifically, the results (Fig. 2-D) suggest participants in the fluent-competent condition rated that their expectations were met ( $M = 4.00, SD = 1.732$ ) significantly higher than participants in the disfluent-incompetent condition ( $M = 1.73, SD = 1.534$ ), ( $p = .001$ ). Additionally, participants in the disfluent-competent condition rated that their expectations were met ( $M = 3.47, SD = 1.356$ ) significantly higher than participants in the disfluent-incompetent conditions ( $M = 1.73, SD = 1.534$ ), ( $p = .019$ ). **Discussion:** The results indicate that participants in the fluent-competent and disfluent-competent conditions perceived the robot to have met their expectations more than the participants in the disfluent-incompetent condition did. Additionally, there was no significant difference found in the rating for the fluent-incompetent and disfluent-incompetent conditions. This result may suggest that task competency has a greater impact on the perception of the robot meeting expectations than language fluency.

#### *RQ3: Effects of Varying Robot Conditions on Perception of Robot as a Teammate*

Separate univariate ANOVAs on the dependent variables revealed significant differences between the varying robot conditions on ratings of the robot as a teammate ( $F(3, 56) = 7.188, p < .0005, \eta^2 = .278$ ). Specifically, the results (Fig. 2-E) suggest participants in the fluent-competent condition rated the robot as a teammate ( $M = 4.13, SD = 1.356$ ) higher than participants in the disfluent-incompetent condition did ( $M = 1.87, SD = 1.552$ ), ( $p < .001$ ). Additionally, participants in the disfluent-competent condition ( $M = 3.40, SD = 1.242$ ) rated the robot as a teammate higher than participants in the disfluent-incompetent condition did ( $M = 1.87, SD = 1.552$ ), ( $p = .019$ ).

**Discussion:** Participants in the fluent-competent and disfluent-competent conditions rated the robot more highly for being a good teammate than participants in the disfluent-incompetent condition did. Additionally, there was no significant difference in the rating for the fluent-incompetent and disfluent-incompetent conditions. The results may indicate that task competency has a greater impact on the perception of the robot as a teammate than language fluency.

#### *RQ4: Effects of Varying Robot Conditions on Participants' Willingness to Work with Robot Again*

There was no significance of the effects of varying robot conditions on participants' willingness to work with the robot again ( $F(3, 56) = 1.927, p = .136, \eta^2 = .094$ ) (Fig. 2-F).

**Discussion:** The results may suggest that neither language fluency nor task competency impacts whether participants would be willing to work with the robot again.

## VII. OVERALL FINDINGS

Overall, the fluent-competent robot was rated higher than the disfluent-incompetent robot in every perception category except willingness to work again. For **RQ1**, we observed a significant interaction between varying robot conditions on perceptions of the robot's verbal competence, intelligence, and reliability. One of the important findings is that language fluency plays a greater role than task competency on the impact of the robot's perceived verbal competence. Additionally, for the perceptions of intelligence and reliability, both language fluency and task competency have an impact on the participants' perceptions.

For **RQ2** and **RQ3**, we explored the effect of varying robot conditions on ratings of the robot meeting expectations and the robot as a teammate. We discerned a significant interaction between varying robot conditions on perceptions of the robot meeting expectations and on the rating of the robot as a teammate. The results indicate that both the competent robots (fluent and disfluent) were perceived as better at meeting expectations and to be a good teammate than the disfluent-incompetent robot. Therefore, we can reason that task competency may be a factor of greater influence on the perceptions of the robot meeting expectations and the robot as a teammate.

For **RQ4**, there were no significant differences between the effects of varying robot conditions on ratings of willingness to work with the robot again, which may indicate that neither language fluency nor task competency alone had a significant impact on the participant's rating of willingness to work with the robot again. We posit that this may be due to the tasks being relatively low stakes in their impact on participants. For example, the robot failing to perform the task correctly may not have been consequential enough in this context for the participant not to want to work with the robot again. There is a need for further exploration of the confounding factors that may impact the willingness of participants to work with the robot again.

## VIII. CONCLUSIONS

In this paper, we presented a human-robot interaction study to investigate the effects of varying language fluency and task competency on the perception of robots. The results highlight that while language fluency may play a greater role than task competency in the perception of verbal competence, both language fluency and task competency contribute to the perception of intelligence and reliability of the robot. The results also indicate that task competency may play a greater role than language fluency in the perception of meeting expectations and being a good teammate. The findings of this work underscore the importance of language fluency and task competency in the context of social HRI and will enable the development of more intelligent robots in the future. In the future, we plan to examine how the perception of robots in the varying language fluency and task competency conditions from the perspective of multilingual individuals differs from that of the perspective of monolingual individuals.

## REFERENCES

- [1] A. A. Bahishti, "Humanoid robots and human society," *Advanced Journal of Social Science*, vol. 1, no. 1, p. 60–63, Nov. 2017.
- [2] M. S. Yasar and T. Iqbal, "A scalable approach to predict multiagent motion for human-robot collaboration," in *IEEE Robotics and Automation Letters (RA-L)*, 2021.
- [3] M. M. Islam and T. Iqbal, "Multi-gat: A graphical attention-based hierarchical multimodal representation learning approach for human activity recognition," in *IEEE Robotics and Automation Letters (RA-L)*, 2021.
- [4] V. V. Unhelkar, S. Li, and J. A. Shah, "Semi-supervised learning of decision-making models for human-robot collaboration," in *Conference on Robot Learning*, 2020.
- [5] M. M. Islam, M. S. Yasar, and T. Iqbal, "MAVEN: A memory augmented recurrent approach for multimodal fusion," in *IEEE Transaction on Multimedia*, 2022.
- [6] M. M. Islam, R. M. Mirzaiee, A. Gladstone, H. N. Green, and T. Iqbal, "CAESAR: A multimodal simulator for generating embodied relationship grounding dataset," in *NeurIPS*, 2022.
- [7] T. Kanda, H. Ishiguro, M. Imai, and T. Ono, "Development and evaluation of interactive humanoid robots," *Proceedings of the IEEE*, vol. 92, no. 11, pp. 1839–1850, 2004.
- [8] M. M. Islam, A. Gladstone, R. Islam, and T. Iqbal, "Eqqa-mx: Embodied question answering using multimodal expression," in *The Twelfth International Conference on Learning Representations (ICLR)*, 2024.
- [9] M. S. Yasar and T. Iqbal, "Vader: Vector-quantized generative adversarial network for motion prediction," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023, pp. 3827–3834.
- [10] M. de Graaf, "Living with robots: investigating the user acceptance of social robots in domestic environments," Ph.D. dissertation, University of Twente, June 2015.
- [11] M. S. Yasar\*, M. M. Islam\*, and T. Iqbal, "IMPRINT: Interactional dynamics-aware motion prediction in teams using multimodal context," in *ACM Transactions on Human-Robot Interaction*, 2023.
- [12] M. M. Islam, A. Gladstone, and T. Iqbal, "PATRON: Perspective-aware multitask model for referring expression grounding using embodied multimodal cues," in *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI)*, 2023.
- [13] M. S. Yasar and T. Iqbal, "Robots that can anticipate and learn in human-robot teams," in *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2022.
- [14] Y. Wang and J. E. Young, "2022 17th acm/ieee international conference on human-robot interaction (hri)," 2014.
- [15] T. Iqbal, S. Rack, and L. D. Riek, "Movement coordination in human-robot teams: A dynamical systems approach," *IEEE Transactions on Robotics*, vol. 32, no. 4, pp. 909–919, 2016.
- [16] T. Iqbal and L. D. Riek, "Temporal anticipation and adaptation methods for fluent human-robot teaming," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 3736–3743.
- [17] S. M. Anzalone, S. Boucenna, S. Ivaldi, and M. Chetouani, "Evaluating the engagement with social robots," *International Journal of Social Robotics*, vol. 7, pp. 465–478, 2015.
- [18] I. Leite, C. Martinho, and Paiva, "Evaluating the engagement with social robots," *International Journal of Social Robotics*, vol. 5, pp. 291–308, 2013.
- [19] F. Eyssel, D. Kuchenbrandt, F. Hegel, and L. de Ruiter, "Activating elicited agent knowledge: How robot and user features shape the perception of social robots," in *The 21st IEEE International Symposium on Robot and Human Interactive Communication*, 2012, pp. 851–857.
- [20] M. S. Yasar, M. M. Islam, and T. Iqbal, "Posetron: Enabling close-proximity human-robot collaboration through multi-human motion prediction," in *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2024, pp. 830–839.
- [21] S. Nikolaidis, M. Kwon, J. Forlizzi, and S. Srinivasa, "Planning with verbal communication for human-robot collaboration," vol. 7, no. 3, 2018.
- [22] A. Marin Vargas, L. Cominelli, F. Dell'Orletta, and E. P. Scilingo, "Verbal communication in robotics: A study on salient terms, research fields and trends in the last decades based on a computational linguistic analysis," *Frontiers in Computer Science*, vol. 2, 2021.
- [23] J. Woo, J. Botzheim, and N. Kubota, "Conversation system for natural communication with robot partner," in *2014 10th France-Japan/ 8th Europe-Asia Congress on Mecatronics*, 2014, pp. 349–354.
- [24] H. N. Green, M. M. Islam, S. Ali, and T. Iqbal, "ispys a humorous robot: Evaluating the perceptions of humor types in a robot partner," in *AAAI Spring Symposium on Putting AI in the Critical Loop: Assured Trust and Autonomy in Human-Machine Teams*, 2022.
- [25] S. Zhao, "Humanoid social robots as a medium of communication," *New Media & Society*, vol. 8, no. 3, pp. 401–419, 2006.
- [26] H. N. Green, M. M. Islam, S. Ali, and T. Iqbal, "Who's laughing nao? examining perceptions of failure in a humorous robot partner," in *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2022, p. 313–322.
- [27] C. Breazeal, K. Dautenhahn, and T. Kanda, "Social robotics," *Springer handbook of robotics*, pp. 1935–1972, 2016.
- [28] E. C. Grigore, A. Pereira, I. Zhou, D. Wang, and B. Scassellati, "Talk to me: Verbal communication improves perceptions of friendship and social presence in human-robot interaction," in *16th International Conference Intelligent Virtual Agents*. Springer, 2016, pp. 51–63.
- [29] A. Gluszek and J. F. Dovidio, "Speaking with a nonnative accent: Perceptions of bias, communication difficulties, and belonging in the united states."
- [30] A. Pavlenko and A. Blackledge, *Negotiation of identities in multilingual contexts*. Multilingual Matters, 2004, vol. 45.
- [31] O. Kang, D. Rubin, and S. Lindemann, "Mitigating us undergraduates' attitudes toward international teaching assistants," *Tesol Quarterly*, vol. 49, no. 4, pp. 681–706, 2015.
- [32] S. van Waveren, E. J. Carter, and I. Leite, "Take one for the team: The effects of error severity in collaborative tasks with social robots," in *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, 2019, pp. 151–158.
- [33] N. Walker, C. Mavrogiannis, S. Srinivasa, and M. Cakmak, "Influencing behavioral attributions to robot motion during task execution," in *Proceedings of the 5th Conference on Robot Learning*, vol. 164, 08–11 Nov 2022, pp. 169–179.
- [34] M. Häring, D. Kuchenbrandt, and E. André, "Would you like to play with me? how robots' group membership and task features influence human-robot interaction," in *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction*, 2014, p. 9–16.
- [35] B. Kim, K. S. Haring, H. J. Schellin, T. N. Oberley, K. M. Patterson, E. Phillips, E. J. de Visser, and C. C. Tossell, "How early task success affects attitudes toward social robots," in *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 2020, p. 287–289.
- [36] M. F. Jung, J. J. Lee, N. DePalma, S. O. Adalgeirsson, P. J. Hinds, and C. Breazeal, "Engaging robots: Easing complex human-robot teamwork using backchanneling," in *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, 2013, p. 1555–1566.
- [37] S. van Waveren, E. J. Carter, and I. Leite, "Take one for the team: The effects of error severity in collaborative tasks with social robots," in *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, 2019, p. 151–158.
- [38] E. J. Carter, L. M. Hiatt, and S. Rosenthal, "You're delaying my task?! the impact of task order and motive on perceptions of a robot," in *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2022, pp. 304–312.
- [39] M. Salem, G. Lakatos, F. Amirabdollahian, and K. Dautenhahn, "Would you trust a (faulty) robot? effects of error, task type and personality on human-robot cooperation and trust," 2015, p. 141–148.
- [40] P. Lennon, "Investigating fluency in efl: A quantitative approach," *Language Learning*, vol. 40, no. 3, pp. 387–417, 1990.
- [41] B. VanPatten and J. Williams, "Skill acquisition theory robert dekeyser," in *Theories in Second Language Acquisition*. Routledge, 2014, pp. 106–124.
- [42] J. Kormos and M. Dénes, "Exploring measures and perceptions of fluency in the speech of second language learners," *System*, vol. 32, no. 2, pp. 145–164, 2004.
- [43] H. R. Bosker, A.-F. Pinget, H. Quené, T. Sanders, and N. H. de Jong, "What makes speech sound fluent? the contributions of pauses, speed and repairs," *Language Testing*, vol. 30, no. 2, pp. 159–175, 2013.
- [44] M. J. White and Y. Li, "Second-language fluency and person perception in china and the united states," *Journal of Language and Social Psychology*, vol. 10, no. 2, pp. 99–113, 1991.
- [45] T. Bongaerts, C. van Summeren, B. Planken, and E. Schils, "Age and ultimate attainment in the pronunciation of a foreign language," *Studies in second language acquisition*, pp. 447–465, 1997.
- [46] A. Hagi-Mohamed, "Perceptions of nonnative english-speaking grad-

- uate teaching assistants: Identity issues, successes, and challenges in the field of tesl/tesol,” *Culminating Projects in English*, 2018.
- [47] D. Barona, “Native and non-native speakers’ perceptions of non-native accents,” *LL Journal*, vol. 3, no. 2, pp. 1–18, 2008.
- [48] M. J. Munro and T. M. Derwing, “The functional load principle in esl pronunciation instruction: An exploratory study,” *System*, vol. 34, no. 4, pp. 520–531, 2006.
- [49] D. L. Rubin and K. A. Smith, “Effects of accent, ethnicity, and lecture topic on undergraduates’ perceptions of nonnative english-speaking teaching assistants,” *International journal of intercultural relations*, vol. 14, no. 3, pp. 337–353, 1990.
- [50] R. W. Todd and P. Pojanapunya, “Implicit attitudes towards native and non-native speaker teachers,” *System*, vol. 37, no. 1, pp. 23–33, 2009.
- [51] A. James and J. Leather, *Second-language speech: structure and process*. Walter de Gruyter, 2011, vol. 13.
- [52] C. Groff, W. Zwaanswijk, A. Wilson, and N. Saab, “Language diversity as resource or as problem? educator discourses and language policy at high schools in the netherlands,” *International Multilingual Research Journal*, pp. 1–19, 2023.
- [53] R. Jackson and T. Williams, “On perceived social and moral agency in natural language capable robots,” 03 2020.
- [54] M. Scheutz, R. Cantrell, and P. Schermerhorn, “Toward humanlike task-based dialogue processing for human robot interaction,” *AI Magazine*, vol. 32, no. 4, pp. 77–84, 2011.
- [55] Y. Bisk, D. Yuret, and D. Marcu, “Natural language communication with robots,” in *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2016, pp. 751–761.
- [56] N. Wang, D. V. Pynadath, S. G. Hill, and A. P. Ground, “Building trust in a human-robot team with automatically generated explanations,” in *Proceedings of the interservice/industry training, simulation and education conference (IITSEC)*, vol. 15315, 2015, pp. 1–12.
- [57] E. Wenger, M. Bronckers, C. Cianfarani, J. Cryan, A. Sha, H. Zheng, and B. Y. Zhao, ““hello, it’s me”: Deep learning-based speech synthesis attacks in the real world,” in *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, 2021, pp. 235–251.
- [58] A. St. Clair and M. Mataric, “How robot verbal feedback can improve team performance in human-robot task collaborations,” in *ACM/IEEE International Conference on Human-Robot Interaction*, 2015, p. 213–220.
- [59] A.-F. Pinget, H. R. Bosker, H. Quené, and N. H. De Jong, “Native speakers’ perceptions of fluency and accent in l2 speech,” *Language Testing*, vol. 31, no. 3, pp. 349–365, 2014.
- [60] N. Gürbüz, “Understanding fluency and disfluency in non-native speakers’ conversational english,” *Educational Sciences: Theory & Practice*, 2017.
- [61] M. Bakti, “Speech disfluencies in simultaneous interpretation,” *De Crom, Dies* (ed.), 2009.
- [62] T. J. Setiyorini, P. Dewi, and E. S. Masykuri, “The grammatical error analysis found in students’ composition,” *Journal Lensa*, vol. 10, no. 2, 2020.
- [63] Y. B. Al-Shujairi and H. Tan, “Grammar errors in the writing of iraqi english language learners,” *International Journal of Education & Literacy Studies*, vol. 5, no. 4, 2017.
- [64] K. Nonukkhethkong, “Grammatical errors analysis of the first year english major students, udon thani rajabhat university,” in *ACLL2013 Conference Proceedings*. IAFOR, 2013.
- [65] Amazon Web Services, Inc, “Amazon polly,” <https://aws.amazon.com/polly/>, 2021, retrieved: June, 2021.