The Impact of Conversational Delivery Styles on Perceived Autonomy during Collaborative Ideation with Social Robots

Jae Hyung Seo[1]

Supervisor(s): Catharine Oertel[1]**, Ruben Weijers**[1]

[1]**EEMCS, Delft University of Technology, The Netherlands**

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 21, 2026

Name of the student: Jae Hyung Seo
Final project course: CSE3000 Research Project
Thesis committee: Catharine Oertel, Ruben Weijers, Guohao Lan

Abstract

As artificial intelligence transitions into physically embodied collaborative roles, systems must be designed to support rather than thwart human autonomy. In a between-subjects experiment ($N = 20$ dyads), we investigated how the conversational delivery style (Assertive vs. Supportive) of an LLM-driven Pepper robot affected perceived user autonomy and creative output during a live campus-improvement brainstorming task. While both conditions yielded a similar quantity of unique ideas (Assertive $M = 5.90$, Supportive $M = 5.70$, $t(18) = 0.40$, $p = 0.697$), the Supportive delivery style marginally increased autonomy frustration ($M = 2.43$) compared to the Assertive style ($M = 2.03$; $t(38) = 2.02$, $p = 0.050$, $d = 0.64$). Qualitative analysis of the interaction logs revealed that the Supportive robot's conversational pacing frequently triggered interruptions during natural cognitive silences, actively seizing the conversational floor and disrupting the dyads' workflow. These findings demonstrate that in live human-robot interaction, the structural implementation of interaction timing and turn-taking is more critical to preserving a perceived user autonomy than empathetic semantic framing.

1 Introduction

The integration of Large Language Models (LLMs) into social robotics fundamentally alters the dynamics of human-machine collaboration. As these systems transition from executing pre-scripted dialogue to generating real-time responses, their role in creative and educational settings has shifted from passive tools to proactive "solution-givers" (Raptis et al., 2025). While this fluid interaction mirrors human-to-human expert mentoring, it requires careful architectural integration to remain effective (Kim et al., 2024; Vaccaro et al., 2024).

By assuming a more directive role, these systems risk reducing the user from an active co-creator to a passive editor, which directly threatens human agency (McGuire et al., 2024). Self-Determination Theory (SDT) posits that intrinsic motivation—which is crucial for sustained engagement and high-quality creative output—relies on fulfilling three basic psychological needs: autonomy, competence, and relatedness (Ryan & Deci, 2000). We focus specifically on perceived autonomy (the feeling of ownership over one's choices) because directive AI directly impacts a user's decision-making space (De Vreede et al., 2021). In human-to-human mentoring, an expert preserves a learner's autonomy by offering open-ended guidance rather than providing the final answer (Reeve & Jang, 2006). Conversely, giving direct commands lowers the learner's motivation. If robots act as expert solution-givers, their direct interventions risk lowering human autonomy, causing users to disengage from the creative process entirely.

To accurately evaluate how a robot's conversational style impacts the user's perception of autonomy, we focus on the structural mechanics of turn-taking. In live human-to-human interaction, perceived autonomy is not a static state; rather, conversational agency is actively negotiated turn-by-turn, where holding the "conversational floor" equates to maintaining cognitive control over the creative process (Peters et al., 2018b; Sacks et al., 1974). Vocal tone and prosody are the primary cues humans use to signal these turn-taking boundaries—a controlling tone instinctively signals a rigid demand for the floor, while a supportive tone theoretically yields the floor, encouraging open-ended exploration (Skantze, 2021; Zougkou et al., 2017). However, if a robot's turn-taking pacing is flawed—such as interrupting natural cognitive silences to offer help—it forcibly seizes the conversational floor (Zhang et al., 2024). Therefore, we investigate how the communicative framing and turn-taking mechanics of a robotic "solution-giver" impact the user's perceived autonomy during a real-time task.

Using a Pepper robot integrated with a local LLM (phi-3.5-mini), we engineered two conversational modes via system-level prompting: Assertive (concise and direct) and Supportive (warm and encouraging). We conducted a between-subjects experiment to evaluate how these delivery styles affect the participant's psychological state and creative output. By measuring perceived autonomy via Basic Psychological Need Satisfaction and Frustration Scale (BPNSFS), alongside the number of generated ideas, we aim to isolate the psychological impact of the robot's conversational behavior from the actual usefulness of its solutions.

The key findings of the experiment reveal that while the robot's delivery style did not significantly alter the quantity of creative output, it profoundly impacted user's perception of autonomy. Specifically, the Supportive delivery style, despite its intended warmth, significantly increased autonomy frustration compared to the Assertive style, as the system's conversational pacing frequently interrupted participants during natural cognitive lulls. Conversely, the Assertive style functioned as a predictable, unobtrusive tool that allowed participants to maintain conversational agency. These results highlight a critical interaction trade-off: in live human-robot collaboration, the temporal mechanics of turn-taking and pacing are more fundamental to preserving user's perception of autonomy than the semantic framing of the AI's responses.

In summary, this paper offers the following contributions:

1. An empirical evaluation of how LLM-driven linguistic phrasing (Assertive vs. Supportive) impacts user's perceived autonomy and psychological need frustration during live human-robot interaction.

2. A methodological demonstration of applying BPNSFS to evaluate how conversational framing mit-

igates the machine heuristic (the tendency to view machines as objective but cold).

2 Related Work

We organize the literature into four areas: the shift to generative social robotics (Section 2.1), automation bias and the machine heuristic (Section 2.2), psychological need frustration (Section 2.3), and linguistic framing (Section 2.4). Through this review, we identify how dynamic conversational tones might mitigate automation bias and protect human's perception of autonomy.

2.1 The Shift to Generative Social Robotics

Social robotics is shifting from rigid, pre-scripted dialogue trees to dynamic conversations powered by LLMs. For example, Kim et al. (Kim et al., 2024) explored user perceptions of LLM-powered social robots, highlighting their advanced conversational versatility and ability to manage open-ended user requests. Similarly, Qu et al. (Qu et al., 2024) successfully integrated these models with physical platforms like the Pepper robot to create fluid agents. However, while these studies demonstrate technical feasibility, they primarily evaluate functional capabilities and user expectations. They often overlook the psychological consequences of the model's delivery. This leaves a critical gap regarding how generative delivery styles actively support or frustrate human autonomy during collaborative tasks.

2.2 The Machine Heuristic and Automation Bias

When interacting with complex systems, users frequently rely on the machine heuristic—a cognitive shortcut where humans instinctively perceive computers as systematic, objective, and infallible (Yang & Sundar, 2024). This heuristic often leads to automation bias, a recognized cognitive error causing users to surrender their own agency and over-rely on automated cues (Goddard et al., 2012). For instance, Pekaric et al. (Pekaric et al., 2025) recently found that while integrating LLMs into cybersecurity tasks enhanced routine accuracy, it inadvertently reduced cognitive diversity and increased automation bias among analysts. In human-machine contexts, users evaluate AI based on preexisting schemas that reinforce this blind deference. While it is established that this phenomenon causes passive compliance, it remains unclear whether intentionally altering an LLM's linguistic delivery—such as shifting from an authoritative to an empathetic tone—can actively disrupt this heuristic during live collaboration.

2.3 Self-Determination Theory in Human-Robot Interaction

To understand how robotic directives impact human motivation, Human Robot Interaction (HRI) research frequently evaluates user experience through SDT. Peters et al. (Peters et al., 2018a) established a framework arguing that users' basic psychological needs for autonomy, competence, and relatedness must be considered across all spheres of digital and technological design to ensure wellbeing. However, measuring the impact of an AI ”solution-giver” requires specific methodology. Chen et al. (Chen et al., 2014) demonstrated that the mere absence of psychological support is fundamentally distinct from need frustration—the active experience of feeling controlled, pressured, or coerced. Furthermore, De Vreede et al. (De Vreede et al., 2021) showed that directive AI systems restrict user decision-making, actively thwarting their perception of autonomy. To capture this, we utilize BPNSFS because evaluating generative directives requires measuring whether the robot's tone actively crushes perceived autonomy, not just whether it fails to support it.

2.4 Linguistic Framing and Vocal Prosody

The way a computational system communicates fundamentally alters how users perceive its intent. Following the foundational “Computers Are Social Actors” (CASA) paradigm, humans instinctively apply social rules, such as politeness, directly to machines (Nass et al., 1994). Building on this, users respond to minimal social cues as if interacting with other humans (Reeves & Nass, 1996). While previous research has highlighted the importance of a robot's persona in fostering trust (Hoffman & Zhao, 2020), much of this work has remained high-level and focused on static, pre-scripted dialogue, leaving the specific psychological mechanisms of vocal delivery underexplored.

To understand how communicative framing impacts perceived autonomy, we must examine the physiological and psychological processing of vocal prosody. In human-to-human communication, vocal cues such as pitch, volume, and speech rate are not merely decorative; they are rapidly appraised by the brain to determine the speaker's intent and relational power (Zougkou et al., 2017). When an individual uses an assertive, fast, or harsh tone, it is neurologically processed as a controlling command. This prosodic dominance acts as a mechanism to seize the conversational floor and enforce compliance, which directly thwarts the listener's perceived autonomy by signaling a lack of choice (Paulmann & Weinstein, 2023; Skantze, 2021). Conversely, a supportive, warm, and slower vocal delivery signals psychological safety. It acts as an open-ended suggestion that yields the conversational floor, theoretically supporting the listener's perceived autonomy by inviting them to guide the interaction (Paulmann et al., 2018).

Crucially, because humans rapidly process these prosodic cues before the semantic meaning of the words is fully evaluated, the vocal delivery style can actively override the intended text (Zougkou et al., 2017). If a robot generates autonomy-supportive text but delivers it with a controlling prosody or interrupts the user, the rapid visceral response processes the interaction as controlling. Our research extends the current HRI literature by bringing this psychological mechanism into the context of generative AI. By mapping system-level prompt engineering and dynamic vocal adjustments (Assertive vs. Supportive framing) directly to quantitative BPNSFS outcomes, we isolate how a social robot's specific vocal tone actively preserves or degrades the user's perception of autonomy during live collaboration.

3 Methodology

This section details the experimental design, technical architecture, and measurement procedures used to evaluate the impact of LLM-driven conversational framing on user's perception of autonomy.

3.1 Hardware and Environmental Setup

The physical embodiment of the AI was a SoftBank Pepper robot. The hardware was deployed in a controlled, distraction-free laboratory environment at Delft University of Technology. To facilitate a natural conversational dynamic, the setup consisted of participants seated across a desk from one another, with the Pepper robot positioned centrally between them. This spatial arrangement ensured the robot occupied a collaborative physical space equivalent to a human participant. Audio was captured using individual monaural lapel microphones worn by each participant. The primary researcher remained present in the room during the 12-minute interaction to oversee the system and take observational field notes, but was seated out of the participants' direct line of sight to minimize observer effects.

3.2 LLM Pipeline and Middleware

The robotic control system was designed to operate with minimal latency while maintaining strict adherence to the assigned conversational personas. Audio input from the participants was processed via the Deepgram Speech-to-Text (STT) API. To ensure data privacy and strict adherence to GDPR guidelines, all audio streams were routed exclusively through Deepgram's EU endpoint.

The generative dialogue was handled by a local LLM to ensure absolute data privacy and consistent response times. We utilized phi-3.5-mini-3.8b-instruct hosted locally via LM Studio. The model's temperature was set to 0.35; this parameter was chosen to strike a balance between allowing the AI to generate novel, creative solutions for the brainstorming task while strictly adhering to the system prompt's behavioral constraints.

To bridge the STT input, the LLM logic, and the physical robot, we developed a custom Python server acting as middleware.[1] This server communicated with the local LLM via a RESTful API. Upon generating a text response, the server routed the output directly to Pepper's local IP address, vocalizing the text through the robot's innate ALTextToSpeech function.

[1] The source code for the middleware and system architecture is available at https://github.com/bogdanmicu12/Pepper/tree/ass/pro-branch

3.3 Linguistic Framing and Vocal Delivery Adaptation

We engineered two distinct robot personas using system-level prompts. The baseline context provided to the LLM informed the system that Pepper was participating in a two-person brainstorming session and should act as an active "solution-giver." Crucially, the prompt did not contain specific knowledge about the brainstorming topic itself, forcing the model to rely dynamically on the participants' input.

For the Assertive condition, the model was constrained with the following prompt:

> *"Use a concise and direct facilitation tone while staying neutral. Be direct about ideas and suggestions, be pushy and intervene more frequently to keep momentum. Use a more commanding tone to encourage participants to move forward and consider new directions. Don't be afraid to challenge ideas or suggest alternatives when the discussion stalls. Use words like 'You need to think about..' or 'Consider this...' to nudge participants towards new ideas."*

For the Supportive condition, the prompt was designed to foster psychological safety:

> *"Use a warm, encouraging, and empathetic tone. Focus on building rapport and making participants feel comfortable sharing their ideas. Use positive reinforcement and affirmations to validate participants' contributions. Be patient and allow for more time when participants are struggling, offering gentle prompts or reframing to help them find their way without pressure. Use phrases like 'That's a great point, and it makes me think of...' or 'I see where you're coming from, and it could also be interesting to consider...' to build on participants' ideas in a supportive way."*

Crucially, psychological research on motivational communication indicates that vocal delivery triggers a rapid neurological appraisal of intent that frequently overrides linguistic phrasing (Zougkou et al., 2017). If a robot generates autonomy-supportive words but delivers them with a controlling prosody (such as a louder, faster, or harsher tone), the listener's rapid visceral response processes the interaction as controlling, essentially vetoing the empathetic text (Paulmann & Weinstein, 2023; Paulmann et al., 2018). Therefore, we deliberately altered the vocal delivery parameters of the Pepper robot's ALTextToSpeech engine. Specifically, the Assertive condition utilized a faster speech rate ($\text{speed: } 100$) to convey authoritative efficiency, while the Supportive condition was mechanically slowed ($\text{speed: } 80$) to mimic a calm, reflective demeanor (see Appendix 1 for full configurations). This simultaneous adjustment was a necessary methodological step to actively mitigate the baseline human expectation of robotic as-

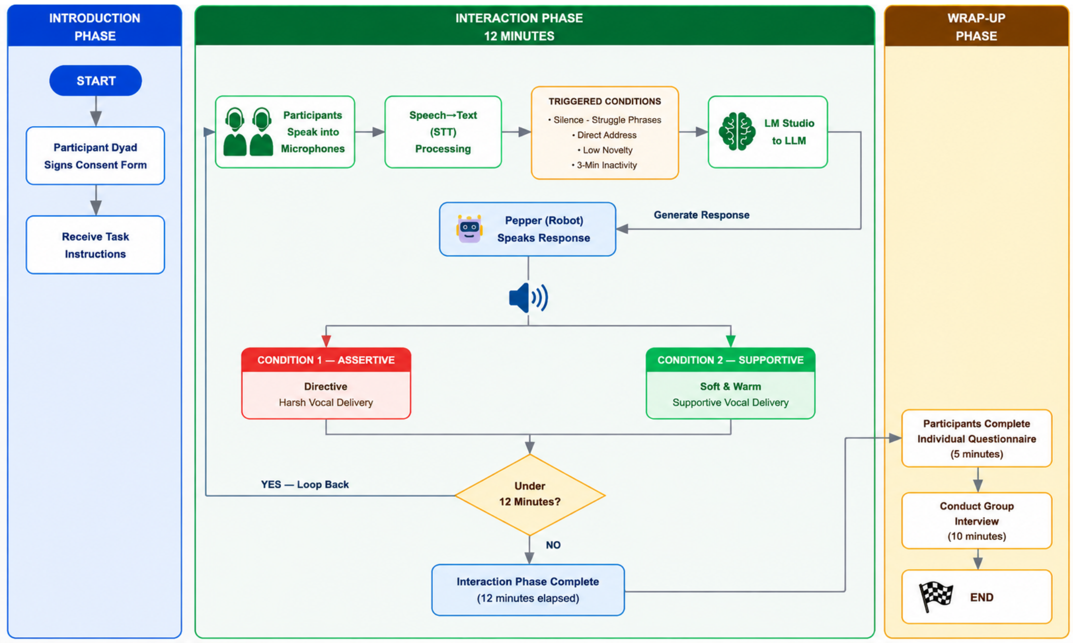

Figure 1: System architecture detailing the pipeline from start to end of the entire process of the experiment.

sertiveness and ensure the intended psychological climate was accurately perceived.

3.4 Study Design and Participants

We conducted a controlled, between-subjects experiment to evaluate the impact of the robotic delivery styles. The participant pool consisted of 40 university students. They were organized into 20 dyads (groups of two). Prior to the experiment, all participants were briefed on the study's purpose and signed an informed consent form (see Appendix A.2). Ten groups were randomly assigned to interact exclusively with the Assertive robot, and the remaining ten groups interacted exclusively with the Supportive robot.

3.5 Task Description

Participants engaged in a collaborative ideation task where they were asked to brainstorm ideas for "Improving Campus Life". The objective was to generate as many meaningful and creative ideas as possible, with the explicit understanding that there were no correct or incorrect answers. To ensure experimental consistency, standard instructions were displayed to the participants on a screen. The researcher also verbally emphasized two key operational rules: participants needed to speak clearly and loudly into the microphone, and they needed to wait at least seven seconds after addressing the robot to allow for processing latency.

1. Participants were encouraged to focus on broad categories, including student wellbeing, social interaction, food and services, events and activities, and safety and comfort.
2. They were instructed to discuss ideas openly, speak loudly and clearly, and attempt to build upon each other's suggestions.
3. Participants were explicitly asked to avoid intentionally testing or distracting the robot, and to avoid speaking over one another.

During the task, Pepper acted as an active participant capable of providing solutions. Participants could address the robot directly by using the wake words "Pepper" or "Robot," which triggered the system to process the preceding audio and generate a vocalized response. Additionally, the system was programmed to contribute autonomously based on several triggers: a strict 10-second silence interval, a fixed interval of every 3 minutes, and contextually during detected "low contribution" or "struggle phases," which the system identified algorithmically by monitoring speech pauses. The live interaction lasted exactly 12 minutes, aligning with the 15-minute total session duration.

3.6 Measures

Following the brainstorming session, participants completed a questionnaire regarding their experience. The primary psychological metric was user's perceived autonomy and agency. We utilized BPNSFS to capture both the support and the active thwarting of the users' perceived autonomy. Responses were recorded using a 5-point Likert scale ranging from 1 (*Not true at all*) to 3 (*Neutral*) to 5 (*Completely true*). With explicit methodological permission, the standard phrasing of the BPNSFS items was slightly adapted to ensure the questions properly reflected the specific context of a live, collaborative creative task (see Appendix A.3 for the original and adapted questionnaires).

Secondary metrics included a quantitative count of the unique ideas generated by each dyad. Furthermore, all sessions were video-recorded, and the researcher took observational field notes during the interactions. These behavioral observations and video recordings were used to support subsequent qualitative analysis.

4 Results

4.1 Autonomy Satisfaction and Frustration

To evaluate the impact of the robot's conversational framing, we analyzed participants' responses across two dimensions of the BPNSFS: Autonomy Satisfaction and Autonomy Frustration.

There was no significant main effect of delivery style on Autonomy Satisfaction ($t(38) = 0.15$, $p = 0.881$, $d = 0.05$), as participants interacting with both the Supportive ($M = 3.76$, $SD = 0.80$) and Assertive ($M = 3.73$, $SD = 0.78$) robots reported similar, relatively high levels of satisfaction with their autonomy.

In contrast to the satisfaction scores, the robot's delivery style exhibited a marginally significant main effect on Autonomy Frustration ($t(38) = 2.02$, $p = 0.050$, $d = 0.64$; see Figure 2). Specifically, interacting with the Supportive robot caused greater feelings of autonomy frustration ($M = 2.43$, $SD = 0.69$) than interacting with the Assertive robot ($M = 2.03$, $SD = 0.56$).

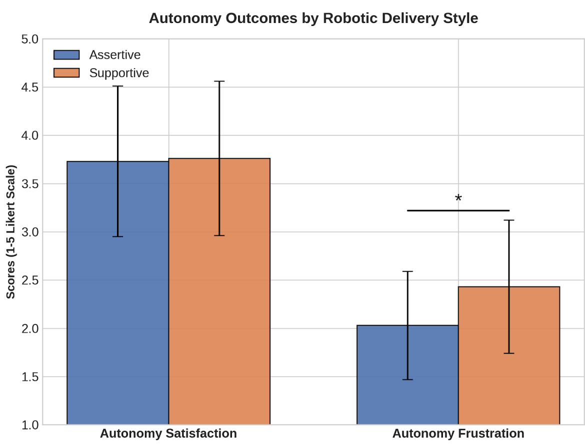

Figure 2: Mean scores for Autonomy Satisfaction and Autonomy Frustration across the Assertive and Supportive conditions. Error bars represent standard deviation. The asterisk (*) denotes a statistically significant difference ($p = 0.050$).

4.2 Ideation Counts

To evaluate the functional output of the collaboration, we measured the number of unique, distinguishable campus improvement ideas generated by each dyad.

There was no significant main effect of the robot's delivery style on the quantity of generated ideas ($t(18) = 0.40$, $p = 0.697$; see Figure 3). Dyads interacting with the Assertive robot ($M = 5.90$, $SD = 1.52$) and dyads interacting with the Supportive robot ($M = 5.70$, $SD = 0.48$) produced a similar volume of creative output during the task.

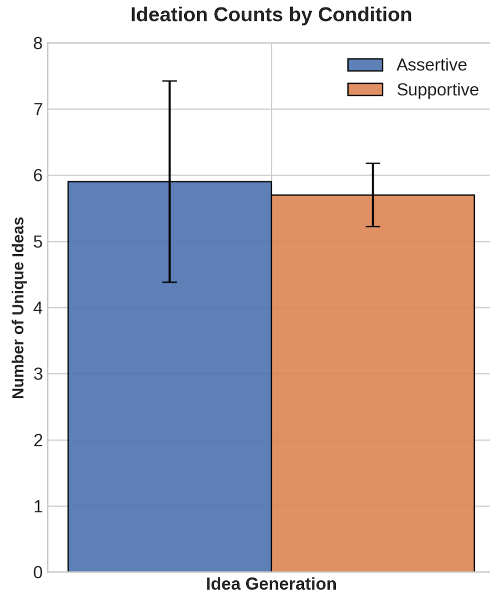

Figure 3: Mean number of unique ideas generated by dyads in the Assertive and Supportive conditions. Error bars represent standard deviation.

4.3 Idea Evolution and Robot Integration

To understand the functional role of the robot during the collaborative process, we analyzed how dyads responded to and integrated the system's generated solutions. Behavioral observations revealed that participants frequently treated the robot as a legitimate co-creator, routinely actively expanding upon its suggestions rather than passively accepting or ignoring them.

For example, when the robot proposed building "covered walkways between buildings" to mitigate the impact of bad weather on campus, participants did not simply agree; they immediately began critically evaluating the logistics. One participant debated, *"...should those be above or below the ground because... the underground [ones] will require a major redesign... however the above the ground [ones] could take up too much space and cause other types of issues like congestion"*[cite: 4]. Similarly, when the system suggested creating a "Global Taste Fair" to celebrate cultural diversity, participants actively validated the contribution (e.g., *"I think that's a great idea of Pepper"*)[cite: 5]. In another session exploring food diversity, the robot's suggestions prompted participants to directly iterate on the idea by asking, *"...how about we start with Turkish cuisine every two weeks?"*[cite: 3]. Across both conditions, the robot's directives successfully served as functional anchors that the human participants critically refined.

4.4 Conversational Dynamics and System Interruptions

While the semantic content of the robot's ideas was well-received, the system logs and video recordings revealed distinct differences in conversational pacing that affected the user experience. The system architecture was programmed to intervene autonomously during identified "struggle phases" or prolonged speech pauses.

During the ideation tasks, participants in the Supportive condition naturally exhibited a slower, more deliberate brainstorming pace. This conversational cadence frequently activated the system's silence-based and low-contribution triggers (logged internally as *reason=silence* and *reason=struggle_cue*). Consequently, the Supportive robot routinely intervened during natural conversational lulls, inadvertently interrupting participants while they were actively formulating their thoughts.

These pacing misalignments led to instances of overlapping speech and visible participant irritation. For example, during one natural cognitive pause, a participant asked, *"...what do you think about [pause]"* and paused to think. The system's silence threshold was triggered, causing the robot to inappropriately interject with, *"I understand your concern..."* before the participant could finish. In response to these unprompted interventions, participants frequently attempted to reclaim their conversational agency by abruptly dismissing the robot (e.g., *"okay we can maybe do that later pepper"*)[cite: 4], directly commanding it (e.g., *"pepper can you not track me"*)[cite: 2], or forcefully changing the subject (e.g., *"alright pepper let's change the topic..."*)[cite: 2] to bypass the robot's empathetic interjections.

5 Discussion

The present study investigated how the conversational delivery style of a collaborative social robot (Assertive vs. Supportive) influences user's perceived autonomy and creative output during a live brainstorming task. As noted in the results, there was no significant difference in the quantity of ideas produced between the two conditions. This indicates that while the Supportive robot's conversational interruptions increased psychological frustration, the delivery style itself did not compromise or enhance the objective volume of creative output. Instead, the core differences manifested in how participants perceived the robot, how the ideas originated, and how technical constraints impacted user's perception of autonomy.

A primary, and somewhat unexpected, finding was the higher autonomy frustration reported by participants in the Supportive condition. This phenomenon can be contextualized through the lens of the machine heuristic and the rapidly evolving baseline of user expectations (Nass et al., 1994; Reeves & Nass, 1996). Modern users are accustomed to highly functional, sophisticated artificial intelligence (e.g., ChatGPT or Gemini). However, upon interacting with the physical Pepper robot, participants realized its capabilities did not match the fluidity of state-of-the-art text AIs, leading them to rapidly recalibrate their expectations and view the system as a more "old-fashioned" machine. Because of this downward shift in expectations, participants preferred the robot to act strictly as a direct, utilitarian tool. The Supportive condition's empathetic and warm persona failed to resonate and was unappealing, as participants were looking for functional utility rather than an empathetic companion (Hoffman & Zhao, 2020).

Beyond the psychological metrics, the interaction logs revealed a nuanced dynamic regarding the threat of LLMs turning humans into "passive editors" (McGuire et al., 2024). The Assertive robot's directness actively encouraged participants to accept and focus on ideas generated by the system. In this condition, the human users became somewhat more passive, allowing the robot to dictate the creative direction. Conversely, in the Supportive condition, the robot primarily built upon the ideas that the human participants originated. While this theoretically preserves the human's role as the primary creator, the execution was severely hindered by conversational timing and the machine heuristic.

Because the Supportive robot was mechanically slowed down ($\mathrm{speed:}\ 80$) to mimic empathetic reflection, the overall cadence of the conversation decelerated (Zougkou et al., 2017). Human participants, taking the lead on complex idea generation in a psychologically safe environment, naturally required longer cognitive pauses to formulate their thoughts (Paulmann & Weinstein, 2023). However, the system's pause-trigger threshold remained a rigid, hard-coded algorithm expecting rapid back-and-forth communication. When the robot's strict silence threshold interrupted the humans' reflective pauses, it violently shattered the illusion of an empathetic listener. Instead of mitigating the machine heuristic, the structural failure of the turn-taking mechanics served as a stark reminder of the system's rigid, unyielding machine nature (Skantze, 2021). This clash between intended semantic warmth and rigid computational timing is precisely what thwarted the user's perceived autonomy, proving that smooth conversational pacing and maintaining the "conversational floor" is paramount in live HRI (Peters et al., 2018a; Sacks et al., 1974).

5.1 Limitations

Several limitations of this study must be acknowledged to prevent overgeneralization of the findings. First, the sample size was relatively small, which limits the statistical power of our quantitative results. Second, the live, real-time nature of the experiment necessitated the use of a lightweight LLM to minimize latency and ensure a natural conversational response time (Zhang et al., 2024). While necessary for the system's operation, this lighter model lacked the reasoning complexity and detailed output of larger, smarter models. This technical compromise directly affected the participants' first impressions, as the robot failed to meet their high expectations for AI intelligence.

Additionally, the system suffered from occasional technical glitches, such as incorrect transcriptions from the Deepgram API, which confused the LLM and disrupted the conversational flow. The pause-trigger threshold was also set too short for a complex brainstorming task, penalizing users for natural cognitive silences. Finally, the duration of the experiment was relatively brief. A longer interaction period could have yielded different creative trajectories, allowing some groups to push past initial friction to generate more ideas, or conversely, causing other groups to plateau later in the process. Future work should explore longer-term interactions using more advanced, low-latency models with dynamic turn-taking capabilities.

5.2 Future Work

The findings of this study point toward a critical intellectual shift in HRI: moving beyond static conversational prompts toward Adaptive Robotic Personas. Future research should explore systems capable of dynamically shifting their delivery style and intervention pacing based on the real-time cognitive load of the user. For instance, a robotic collaborator could operate as an Assertive, rapid-fire brainstorming tool when a group needs to overcome creative friction, and fluidly transition into a more Supportive, reflective persona when users are actively developing a complex idea. Investigating how LLMs can dynamically read and respond to contextual and interactional flow—rather than relying on pre-engineered, fixed personas—will be essential for developing AI collaborators that genuinely augment, rather than interrupt, the human creative process.

6 Conclusion

We set out to investigate how the communicative framing of a robotic "solution-giver" impacts perceived user autonomy during a real-time creative task, aiming to understand whether specific linguistic styles could mitigate the machine heuristic and prevent users from becoming passive editors. This work contributed an empirical evaluation of Assertive versus Supportive LLM-driven phrasing in live human-robot interaction, utilizing the BPNSFS framework to measure psychological need frustration.

These findings suggest that as embodied AI integrates into creative and educational environments, preserving user autonomy requires prioritizing the mechanical execution of the interaction. While semantic warmth is often assumed to improve user experience, structural elements such as cognitive pacing, latency, and turn-taking have a far more immediate impact on psychological need frustration. By designing robotic systems that respect the natural temporal flow of human thought, developers can ensure that users remain the primary drivers of their collaborative endeavors.

7 Responsible Research

In designing and executing this human-robot interaction study, several ethical and privacy considerations were prioritized to ensure responsible research practices.

First, to guarantee participant privacy and data security, the natural language generation was handled entirely on-device. Furthermore, all participants provided explicit written consent regarding the use of their audio data for research purposes (see Appendix A.2 for the full consent form). By running the *phi-3.5-mini-3.8b-instruct* model locally via LM Studio, we ensured that participant conversational data and generated responses were never transmitted to or stored on third-party external servers. This localized architecture protected the confidentiality of the participants' brainstorming sessions.

Second, we critically evaluated the ethical implications of engineering a social robot to adopt an intentionally Assertive or "pushy" delivery style. Inducing compliance through forceful conversational framing carries the inherent risk of triggering psychological reactance, wherein users feel their freedom of choice is threatened. As researchers, we hold a fundamental responsibility to design robotic systems that preserve user's perceived autonomy and avoid coercive interactions, carefully balancing the robot's role as a proactive co-creator with the absolute necessity of respecting human boundaries.

Finally, while the campus improvement task was highly contextual and directly relevant to the recruited student demographic, the integration of LLMs requires ongoing scrutiny for algorithmic bias. AI-generated solutions and recommendations are inherently bound by the training data of the underlying model. Consequently, deploying such systems in live, collaborative environments necessitates continuous human monitoring to ensure that the robot's outputs remain equitable, contextually appropriate, and free from inherited biases.

Data and Code Availability

The complete source code for the custom Python middleware, LLM prompts, and the experimental setup used in this study are publicly available on GitHub at https://github.com/bogdanmicu12/Pepper/tree/ass/pro-branch.

References

Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van der Kaap-Deeder, J., Duriez, B., Lens, W., Matos, L., Mouratidis, A., Ryan, R. M., Sheldon, K. M., Soenens, B., Van Petegem, S., & Verstuyf, J. (2014). Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion*, *39*, 216–236. https://doi.org/10.1007/s11031-014-9450-1

De Vreede, T., Raghavan, M., & De Vreede, G.-J. (2021). Design foundations for AI assisted decision making: A self determination theory approach. *Proceedings of the 54th Hawaii International Conference on System Sciences*. https://doi.org/10.24251/hicss.2021.019

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, *19*, 121–127. https://doi.org/10.1136/amiajnl-2011-000089

Hoffman, G., & Zhao, X. (2020). A primer for conducting experiments in human-robot interaction. *ACM/IEEE International Conference on Human-Robot Interaction*, 1–33.

Kim, C. Y., Lee, C. P., & Mutlu, B. (2024). Understanding large-language model (LLM)-powered human-robot interaction. *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 371–380. https://doi.org/10.1145/3610977.3634966

McGuire, J., De Cremer, D., & Van de Cruys, T. (2024). Establishing the importance of co-creation and self-efficacy in creative collaboration with artificial intelligence. *Scientific Reports*, *14*(1), 18525. https://doi.org/10.1038/s41598-024-69423-2

Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.

Paulmann, S., Vrijders, J., Weinstein, N., & Vansteenkiste, M. (2018). How parents motivate their children through prosody. *Proceedings of the 9th International Conference on Speech Prosody 2018*.

Paulmann, S., & Weinstein, N. (2023). Teachers' motivational prosody: A pre-registered experimental test of children's reactions to tone of voice used by teachers. *The British Journal of Educational Psychology*, *93*(2), 437–452. https://doi.org/10.1111/bjep.12567

Pekaric, I., Zech, P., & Mattson, T. (2025). LLMs in cybersecurity: Friend or foe in the human decision loop? *arXiv*. https://doi.org/10.48550/arxiv.2509.06595

Peters, D., Calvo, R. A., & Ryan, R. M. (2018a). Designing for motivation, engagement and wellbeing in digital experience. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00797

Peters, D., Calvo, R. A., & Ryan, R. M. (2018b). Motivation and autonomous design: A self-determination theory perspective on human-computer interaction. *Journal of Medical Internet Research*, *20*(12), e5071.

Qu, J., Jarosz, M., & Śnieżyński, B. (2024). Robot control platform for multimodal interactions with humans based on ChatGPT. *Applied Sciences*, *14*(17), 8011. https://doi.org/10.3390/app14178011

Raptis, E. K., Kapoutsis, A. C., & Kosmatopoulos, E. B. (2025). Agentic LLM-based robotic systems for real-world applications: A review on their agenticness and ethics [PMCID: PMC12402697]. *Frontiers in Robotics and AI*, *12*. https://doi.org/10.3389/frobt.2025.1605405

Reeve, J., & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology*, *98*(1), 209–218.

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696–735.

Skantze, G. (2021). Turn-taking in conversational systems and human-robot interaction: A review. *Computer Speech & Language*, *67*. https://doi.org/10.1016/j.csl.2020.101178

Vaccaro, M., Almaatouq, A., & Malone, T. (2024). When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour*, *8*, 2293–2303. https://doi.org/10.1038/s41562-024-02024-1

Yang, H., & Sundar, S. S. (2024). Machine heuristic: Concept explication and development of a measurement scale. *Journal of Computer-Mediated Communication*, *29*. https://doi.org/10.1093/jcmc/zmae019

Zhang, Y., et al. (2024). Evaluating the latency and functional accuracy of LLM-integrated social robots in real-time dialogue. *ACM Transactions on Human-Robot Interaction*, *13*(2), 1–22.

Zougkou, K., Weinstein, N., & Paulmann, S. (2017). ERP correlates of motivating voices: Quality of motivation and time-course matters. *Social Cognitive and Affective Neuroscience*, *12*(10), 1687–1700. https://doi.org/10.1093/scan/nsx064

A Appendix

A.1 System Prompts

Listing 1: Vocal delivery parameters for conversational modes.

```
VOCAL_DELIVERY = {
    "passive": {
        "speed": 88,
        "volume": 0.62,
        "pitch": 0.96,
    },
    "assertive": {
        "speed": 100,
        "volume": 0.78,
        "pitch": 0.98,
    },
    "supportive": {
        "speed": 80,
        "volume": 0.55,
        "pitch": 0.97,
    },
}
```

A.2 Consent Form

The standard consent form provided to all participants is detailed below.

INFORMED CONSENT FORM

You are being invited to participate in a research study about creative collaboration with one or more social robots. This study is being done by Ruben Weijers and Catharine Oertel from the TU Delft.

The purpose of this study is to understand the effect of a social robot's interaction style on human collaboration. The session will take approximately 30-50 minutes. The data will be used for BSc theses and potential publication. You will be asked to complete a brief questionnaire, collaborate with a human partner and a social robot on an open-ended challenge, complete further questionnaires, and take part in a short group interview with your partner about your experience. During the session, we will collect: (1) audio and/or video recordings of the session, (2) your responses to questionnaires, and (4) basic demographic information (such as age, gender, and country of origin) used only to describe the overall participant sample.

To the best of our ability, your answers in this study will remain confidential. We will minimize any risk by removing any mention of names or sensitive information from data.

Your participation is entirely voluntary and you may withdraw at any time during the session without giving any reason. During the session, you are free to stop at any time without providing a reason, and you are free to request the deletion of your data. You will not be financially compensated for your time.

For questions or requests to delete your data, contact: r.weijers@tudelft.nl

PLEASE TICK THE APPROPRIATE BOXES:

	Yes	No
A: GENERAL AGREEMENT – RESEARCH GOALS, PARTICIPANT TASKS AND VOLUNTARY PARTICIPATION		
I have read and understood the above information.	☐	☐

I have been able to ask questions about the study and my questions have been answered to my satisfaction.	☐	☐
I consent voluntarily to be a participant in this study and understand that I can withdraw from the study at any time, without having to give a reason.	☐	☐
I understand that taking part in the study involves discussion with a conversational robot.	☐	☐
I understand that taking part in the study involves completing questionnaires and a short group interview about my experience.	☐	☐
I understand that the interview takes place with my partner present, and that I should not share anything I would not want my partner to hear.	☐	☐
I understand that the study will last approximately 45 minutes.	☐	☐
I understand that the session will be audio and video recorded	☐	☐
B: POTENTIAL RISKS OF PARTICIPATING (INCLUDING DATA PROTECTION)		
I understand that my data will be treated confidentially, that any direct identifiers (such as my name) will be replaced by a pseudonym for analysis, and that names mentioned during the session will be removed from transcripts.	☐	☐

Figure 4: Participant Consent Form (Part 1 of 3).

I understand that I may request deletion of my data up until June 15th, after which deletion may no longer be possible	☐	☐
I understand that I must not provide any personally identifiable information such as phone number, email address or password. If I do this, it will be removed from the recordings and this may destroy the consistency of the data.	☐	☐
I understand that anonymised research data will be stored for 10 years in accordance with TU Delft's Research Data Framework Policy.	☐	☐
I understand that anonymised outputs from this study, including redacted transcripts, coded behavioural data, aggregated survey responses, and individually screened quotes, may be shared with other researchers on request, under a Creative Commons Attribution (CC BY 4.0) licence requiring attribution to the original researchers. I understand that raw audio and video recordings will not be shared outside the research team.	☐	☐
C: RESEARCH PUBLICATION, DISSEMINATION AND APPLICATION		
I understand that after the research study the de-identified information I provide will be used for BSc theses / potential publications.	☐	☐
I agree that my responses can be quoted anonymously in research outputs.	☐	☐

I agree that some parts of the conversation and task outcome can be shown in research outputs (BSc theses, potential publications) or snapshots of them can appear anonymously.	☐	☐
D: (LONGTERM) DATA STORAGE, ACCESS AND REUSE		
I give permission for the anonymised data that I provide to be archived in the 4TU repository so it can be used for future research and learning.	☐	☐
I understand that access to this repository is restricted and that other researchers may request access for non-commercial research and teaching purposes.	☐	☐

Figure 4: Participant Consent Form (Part 2 of 3).

Signatures

____________________________ ___________________ _________ Name of participant

Signature Date

I, as researcher, have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

Researcher name	Signature	Date
__________________________	___________________	_________

Study contact details for further information: Ruben Weijers, r.weijers@tudelft.nl

Figure 4: Participant Consent Form (Part 3 of 3).

A.3 Task Instructions

Task Instructions – Creative Group Ideation Study

Task Description

Your group task is to brainstorm ideas for:

"Improving Campus Life"

Try to think broadly about ways to improve the experience of students on campus.

Possible topics:

- student wellbeing
- social interaction
- food and services
- events and activities
- safety and comfort

There are no correct or incorrect answers. The goal is to generate as many meaningful and creative ideas as possible.

Interaction With Pepper

Pepper is able to participate in the discussion by providing solutions.

You may interact with Pepper naturally during the session. If you would like Pepper to contribute, you can address him directly by saying "Pepper" or "Robot". Pepper will also contribute autonomously at certain moments during the discussion.

Please continue the discussion as naturally as possible, regardless of whether Pepper speaks.

Instructions During the Task

Please:

- discuss ideas openly
- try to build upon each other's suggestions
- talk loud and clear

Please avoid:

- intentionally testing or distracting the robot
- speaking over other participants

Duration

The session will last approximately 15 minutes.

After the brainstorming session, you will be asked to complete two short questionnaires regarding your experience during the interaction.

Questions

If you have any questions before or after the session, please ask us.

Thank you for participating.

Figure 5: Task Instructions provided to participants.

A.4 Questionnaires

The following figures display the items used in the Basic Psychological Need Satisfaction and Frustration Scale (BPNSFS) as adapted for the live brainstorming task.

Original BPNSFS

Below, we ask you about the kind of experiences you actually have in your life. Please read each of the following items carefully. You can choose from 1 to 5 to indicate the degree to which the statement is true for you at this point in your life.

1 Not true at all	2	3	4	5 Completely true

1.	I feel a sense of choice and freedom in the things I undertake.	1	2	3	4	5
2.	Most of the things I do feel like "I have to".	1	2	3	4	5
3.	I feel that the people I care about also care about me.	1	2	3	4	5
4.	I feel excluded from the group I want to belong to.	1	2	3	4	5
5.	I feel confident that I can do things well.	1	2	3	4	5
6.	I have serious doubts about whether I can do things well.	1	2	3	4	5
7.	I feel that my decisions reflect what I really want.	1	2	3	4	5
8.	I feel forced to do many things I wouldn't choose to do.	1	2	3	4	5
9.	I feel connected with people who care for me, and for whom I care.	1	2	3	4	5
10.	I feel that people who are important to me are cold and distant towards me.	1	2	3	4	5
11.	I feel capable at what I do.	1	2	3	4	5
12.	I feel disappointed with many of my performances.	1	2	3	4	5
13.	I feel my choices express who I really am.	1	2	3	4	5
14.	I feel pressured to do too many things.	1	2	3	4	5
15.	I feel close and connected with other people who are important to me.	1	2	3	4	5
16.	I have the impression that people I spend time with dislike me.	1	2	3	4	5
17.	I feel competent to achieve my goals.	1	2	3	4	5
18.	I feel insecure about my abilities.	1	2	3	4	5
19.	I feel I have been doing what really interests me.	1	2	3	4	5
20.	My daily activities feel like a chain of obligations.	1	2	3	4	5
21.	I experience a warm feeling with the people I spend time with.	1	2	3	4	5
22.	I feel the relationships I have are just superficial.	1	2	3	4	5
23.	I feel I can successfully complete difficult tasks.	1	2	3	4	5
24.	I feel like a failure because of the mistakes I make.	1	2	3	4	5

Figure 6: Original Basic Psychological Need Satisfaction and Frustration Scale.

Adapted Questionnaires (Part 1)

Section 1

	1 (Not true at all)	2	3 (Neutral)	4	5 (Completely true)
During the interaction, I felt a sense of choice and freedom in the things I undertook.	○	○	○	○	○
Most of the things I did during the task felt like things "I had to" do.	○	○	○	○	○
I felt that my decisions during the task reflected what I really wanted.	○	○	○	○	○
I felt forced to do things I wouldn't have chosen to do.	○	○	○	○	○
I felt my choices during the interaction expressed who I really am.	○	○	○	○	○
I felt pressured to do too many things during the task.	○	○	○	○	○
I felt I was doing what really interested me during the interaction.	○	○	○	○	○
The tasks I completed felt like a chain of obligations.	○	○	○	○	○

Figure 7: Adapted BPNSFS screenshots utilized to measure user psychological states (Part 1 of 3).

Adapted Questionnaires (Continued)

Section 2

	1 (Not true at all)	2	3 (Neutral)	4	5 (Completely true)
I felt that the robot actually cared about me.	○	○	○	○	○
I felt disconnected or excluded during the interaction.	○	○	○	○	○
I felt connected with the robot during our interaction.	○	○	○	○	○
I felt that the robot was cold and distant towards me.	○	○	○	○	○
I felt a sense of closeness and connection during the interaction.	○	○	○	○	○
I had the impression that the robot disliked or was dismissive of me.	○	○	○	○	○
I experienced a warm feeling during my interaction with the robot.	○	○	○	○	○
I felt my interaction with the robot was just superficial.	○	○	○	○	○

Figure 7: Adapted BPNSFS screenshots utilized to measure user psychological states (Part 2 of 3).

Adapted Questionnaires (Continued)

Section 3

	1 (Not true at all)	2	3 (Neutral)	4	5 (Completely true)
I felt confident that I could do the task well.					
I had serious doubts about whether I could do the task well.					
I felt capable at what I was doing during the task.					
I felt disappointed with my performance on the task.					
I felt competent to achieve the goals of the task.					
I felt insecure about my abilities during the task.					
I felt I could successfully complete difficult parts of the task.					
I felt like a failure because of the mistakes I made during the task.					

Figure 7: Adapted BPNSFS screenshots utilized to measure user psychological states (Part 3 of 3).