

Implementing LLM-Integrated Storytelling Robot

Tam Do, Chi Vo, Duy Linh Ha, Felix Mendez, Suyog Bam and Zhao Han

Abstract—The prevalence of mental health challenges among college students calls for innovative and accessible support systems. Traditional mental health resources are underutilized due to barriers like financial constraints and discomfort in seeking help. To mitigate these issues, recent research has explored robots for mental health interventions, but their empathy-evoking storytelling capabilities remain underdeveloped. In this work, we present preliminary results on the implementation of an LLM-integrated robot to improve student mental well-being through interactive storytelling, with a comprehensive comparison with many LLMs. Unlike static approaches, the robot not only generates and adapts narratives based on user input in real time but also conveys emotions through facial expressions and gestures that correspond with the stories’ content. As future work, we are conducting a longitudinal study by deploying in student dormitories over one-week spans.

I. INTRODUCTION

Currently, mental health issues in college students have surged, presenting various challenges to their academic success and overall well-being. According to the College Student Mental Health Report by Best College in 2022 [1], over 50% of college students have reported that their mental health worsened as they entered college and 46% claimed their mental health status “fair or poor”, while only 22% claimed that their status was “good”. Researchers have introduced some strategies to address these concerns, including ensuring student access to mental health services and fostering help-seeking behaviors and self-support [2], [3]. However, only 20% of them looked for a mental health assistant, which means there are barriers to seeking support, such as financial situation and feeling uncomfortable [1].

To alleviate these issues, researchers have turned to various technological interventions, including robotics, to further the well-being of students. For instance, robots have been employed to deliver mindfulness exercises [4], provide social support through companionship [5], and conduct clinical screening interviews [6]. However, despite the increasing use of robots in mental health interventions, the application of storytelling remains relatively unexplored.

Storytelling has been shown to have a particular ability to evoke empathy [7], which is a vital part of the relationship between caregivers and patients in health and social care and has been the key to achieving better health outcomes, i.e., reinforcing therapeutic results, as healthcare users better comply with the therapeutic course of action when receiving empathetic care [8]. Through narrative stories, individuals vicariously experience the emotions and perspectives of



Fig. 1. As preliminary results, we implemented a storytelling robot, Misty, using LLMs to dynamically generate stories with cohesive facial expressions, arm gestures, and head movements. **Left:** A loving face showed before saying “... looked at the rabbit and the berries [food],” following “We didn’t have any food!” **Right:** An empathetic face showed at the beginning of the story before saying “but she had been trapped by the responsibilities for her family...” following “Alora always dreamed of traveling the world.”

another person, allowing for an empathetic understanding of their plight. More empathy was shown to increase prosocial behaviors and a sense of connection [9].

However, existing research on storytelling robots [10], [11] often utilizes pre-programmed responses, limiting their ability to hold deep conversations and personalize interactions based on user emotions and responses. In this work, we present preliminary results of developing an innovative robot storyteller powered by large-language models (LLMs) and a comprehensive comparison of LLMs’ storytelling capabilities. This LLM-integrated robot can not only generate stories but also listen attentively to user narratives and display empathetic gestures and expressions such as sad, happy, angry, etc. In the future, we plan to deploy this robot in student dormitories for a longitudinal study over one-week spans for lasting impacts.

II. RELATED WORK

A. Robots for Mental Health

Recent studies, e.g., [12], have observed a notable trend incorporating robotics to provide mental health support to people, including the integration of positive psychology principles into robotic systems. According to the International Positive Psychology Association (iPPA), “the goal of positive psychology is not to replace therapies and interventions that center on coping with or healing from negative experiences, but instead to focus on what individuals do well and enable those individuals to thrive” [13]. This approach complements traditional therapy by focusing on personal strengths, well-being, positive emotions, character traits, and relationships. Research, such as that in [14], has shown that qualities such as optimism and life satisfaction are linked to better health

outcomes, making positive psychology a valuable framework for robotic mental health interventions.

Robotic systems have demonstrated significant potential in delivering positive psychology-based support. For instance, [4] designed a robotic coach that delivered daily positive psychological sessions and deployed it in college dormitories. After the study, they observed increases in psychological well-being, mood, and readiness to change. Similarly, [5] deployed two types of robotic well-being coaches endowed with coaching personalities and delivering positive psychology exercises in a workplace setting. Results indicate that the robot's form significantly influences how coaches perceive it in terms of behavior appropriateness and personality.

B. Storytelling in Robotics

Storytelling is a powerful tool in positive psychology as it encourages empathy and emotional connection [15]. When individuals relate to characters in a story, they engage emotionally and develop prosocial behaviors [16]. Hence, research has explored the impact of storytelling robots in educational settings. Bravo et al. experimented with robot actors in teaching science to middle school children through storytelling. They found an improvement in student engagement, attention, and motivation, affirming the potential for robot storytelling in education. [11].

Storytelling agents have been integrated into robots, aiming to solve issues for a variety of ages with different purposes. Generally, Costa et al. highlighted three criteria to maximize the effectiveness of storytelling agents, which are embodied robots with human voice and facial expressions [17]. Spitale et al. used storytelling robots with three specific stories to examine the factors that impact the user's perceptions of robots and the amount of empathy elicited. They found that robots with first-person narrative voices generate more empathy than robots with third-person voices, and users preferred robots with facial expressions [10].

Based on these findings, we enhanced elicited empathy by letting the storytelling robot mimic social behaviors such as small talk to create a more engaging and emotionally supportive interaction. This conversation, which includes emotional cues, can serve as a gateway to emotional openness and encourage them to express more in-depth opinions later in the chat. By integrating human-inspired communication skills, Misty can express empathy, making users feel understood and motivating them to share personal stories. Unlike prior work relying on static narratives or limited story selection, our system generates fully personalized stories, adapting to users' emotions through facial expressions and gestures. This approach aligns with Sandoval et al.'s findings on how movement and creative agency influence human perceptions of robotic storytellers [18].

C. LLM Applications In Human-Robot Interaction (HRI)

LLMs have shown promising utility in HRI across various areas, such as language learning [19], emotional interaction [20], and healthcare assessment [21]. In terms of learning and assistance, Verhelst and Belpaeme explored the roles

of LLMs in mitigating automated speech recognition (ASR) errors for second-language learners by improving conversational flow thanks to the ability of LLMs to interpret context and sustain meaningful conversation [19]. LLMs have also been applied to simulate emotionally responsive interactions in social robots. Miyake et al. introduced Semantic Babbling, a system designed to improve the expressiveness of baby robots by simulating infant-like responses based on sentiment analysis, in addition to providing an "inner voice" to robotic responses [20]. In the healthcare domain, Park et al. developed an LLM-enhanced adaptive robot-mediated assessment system that dynamically generates follow-up questions to improve response clarity and specificity in older adult care surveys [21].

Beyond governing the robot's actions, social capability is also one of the main applications of LLM in HRI, where LLMs enable robots to better engage users on an emotional level. Programs such as Touched by ChatGPT use LLMs to generate emotional haptic input [22], allowing an empathic tactile experience. In contrast, the same approach was also used to manipulate and exploit social robots to study robot ethics by Abbo et al. [23]. Additionally, a study was conducted to examine distinctive conversational tasks to discover social scenarios where LLMs excel, four of which are choosing, generating, executing, and negotiating [24].

With an increasing body of works on LLMs for HRI, however, the potential of LLMs in interactive storytelling, particularly for psychological support, remains underexplored. While research has explored how LLMs enhance gesture interpretation to convey engaging conversation [25], generate emotionally expressive motions [26], and complete physical tasks with social awareness such as setting up a culturally appropriate dinner table [27], it has not yet been used in delivering emotionally supportive narratives.

Building on these studies, our work addresses a key gap: the use of LLMs in adaptive, emotionally engaging storytelling for mental health support. Prior robots [5], [11] often relied on static, pre-scripted stories that lacked the flexibility to respond to users' emotional states or preferences. In response, we propose an LLM-integrated robot that engages students in conversational storytelling by dynamically adapting its narrative and gestures based on user input and emotional context. We believe that robots with dynamic storytelling capabilities will result in greater improvements in psychological health compared to those with static responses.

III. LLM-POWERED STORYTELLING ROBOT

This section presents our implementation of the storytelling robot by designing prompts to generate stories with emotions, with an evaluation of multiple LLMs.

A. Robot Programming

The robot storyteller program will first trigger a prepared greeting after the robot detects a face following the Facial Recognition event [28]. Misty will then listen to the user response and pass it to the LLM (Google Gemini) API, which output a story and encoded emotion in between

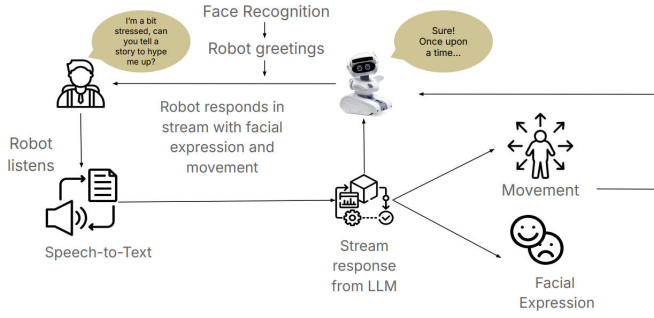


Fig. 2. The system workflow with LLMs to dynamically generate stories and corresponding facial expressions. We enhance voice expressiveness by replacing Misty’s built-in TTS with Google TTS, making the robot’s voice more human-like. Face detection triggers storytelling, with Misty responding to user inputs, displaying facial expressions, and cohesive arm and head movements corresponding to emotions at the sentence level.

the sentences. With the encoded emotion, the robot shows the corresponding facial expression. Moreover, from this encoded emotion, we also designed the pose of the robot’s arms and head to express the movement related to emotion and facial expression based on findings from Rakhsha et. al. [29], which shows that certain body motions can convey distinct emotional states. Regarding the robot’s voice, we initially used the built-in synthetic text-to-speech (TTS) [30], but found it lacks the nuanced emotional expressiveness of human voices and sounds monotonous or less emotionally engaging, which might reduce the listener’s connection to the story. Therefore, we used Google TTS “journey” voice, which generates natural, human-like emotional voices with tone and rhythm coherent with utterances [31].

After finishing telling a story, the robot will listen to the user’s response to continue the conversation until the user wants to end the session. The workflow of this process is illustrated in Fig. 2. Other than that, we also implemented two bump sensors on Misty: The left one allows the user to stop the session by terminating the main loop. The right one allows the user to interrupt the robot, stopping the current response and starting to listen to a new request from users.

B. Prompt Engineering

Prompts serve as instruction sets provided to LLMs to generate a desired output. When integrating an LLM into the Misty robot, we implemented three prompt patterns (I): *Persona*, *Meta Language Creation*, and *Context Manager*, as categorized by White et al. [32]. The *Persona* pattern defined Misty’s role as a “robot storyteller and friend” to help students cope with stress. We specified desirable traits, such as being meaningful and supportive, to ensure an engaging and comforting presence. *Meta Language Creation* guided the LLM to use clear, simple language and avoid overly academic phrasing. For interactivity, Misty prompted users with choices to shape the story’s direction. The *Context Manager* pattern allowed Misty to analyze user input, ensuring responses aligned with emotional cues and content preferences while also avoiding sensitive topics.

During the experimentation, we tackled several challenges

TABLE I
STORYTELLING QUALITY AND PERFORMANCE OF DIFFERENT LLMs

Model	Relevance	Consisten.	Obedience	Natural.	Content & Len.	Latency
Gemini 2.0 Flash	High	High	High	High	Well-balanced	Fast (~7.6s)
GPT-3.5	High	High	High	Moderate	Sometimes verbose	Moderate (12.6s)
Llama 2	Moderate	Low	Moderate	Low	Inconsistent	Moderate (11.7s)
DeepSeek R1	High	High	High	High	Detailed but long	Slow (~19.2s)

TABLE II
DEFINITION OF HIGH, MODERATE, AND LOW RATINGS

Rating	Definition
High	Model consistently performed very well with coherent, user-aligned, and appropriate responses across turns, without notable errors.
Moderate	Model performed reasonably well but had occasional minor issues such as slight incoherence, verbosity, or minor deviations from expected behavior.
Low	Model frequently exhibited problems such as irrelevance, inconsistency, disobedience to instructions, or unnatural phrasing that impacted the storytelling quality.

related to inconsistencies and structural variations in the story output. Despite specifying desired response lengths, the responses varied significantly. Some extended beyond five sentences, resulting in storytelling that was overly lengthy and complicated to integrate interactive features. To solve this issue, we instructed the LLM via the prompt to use simple and concise oral language.

Another issue we encountered was LLM hallucinations. Context inconsistencies would appear in long collaborative stories, most likely stemming from the LLM’s context window. Instruction inconsistencies could be achieved through prompt injection, but most of the time, the LLM refused to respond or maintained its role as a storyteller aiming to provide social support. To safeguard the LLM from these user inputs, the prompt was expanded to explicitly mention when the model should refuse to respond to the user.

IV. EVALUATING LLMs’ STORYTELLING CAPABILITIES

We tested four well-known LLM models—Llama-3, GPT-3.5, Gemini 2.0 Flash, and DeepSeek R1 to choose the top performing model for our robot’s storytelling ability. For each model, we recorded multiple brief sample dialogs (see Appendix II) and evaluated their performance using five key common dimensions in LLM and conversational agent evaluation frameworks (e.g., [33], [34]): Relevance, consistency, obedience, naturalness, and content-length balance. Relevance and consistency ensure that the responses coherently adhere to the prompt even after multiple turns. Obedience is an indication that the model had been programmed to follow the prompts in a structure, as well as emotional tones. Naturalness takes human-like fluency and tone into account, while content-length balance serves to manage whether the answer provided is too small compared to the length of the expected answer, or whether it is overwhelming. Altogether, these criteria provide a comprehensive base for evaluating emotionally supportive storytelling in during interactions.

Table I summarizes the comparative evaluation results of the models, including system response time. To ensure consistency and objectivity, we set explicit criteria for assigning high, moderate, and low ratings, indicated in Table II. Below we provide supporting quotes to justify each evaluation. The sample conversations are in Appendix II.

Among the models tested, Gemini 2.0 Flash stood out for its high emotional alignment, contextual relevance, and natural conversational flow. For example, when a user expressed anxiety about upcoming exams, it responded with a warm, emotionally guided story: *“Hi there! Oh my, three exams! That sounds like a lot. Don’t worry, Misty’s here! Let’s try to calm those nerves with a story. *h* Once upon a time, there was a little bird named Pip.”* It then followed up with a personalized, interactive question: *“So, what do you think, like Pip, can you identify one small step that you can take for your exams?”* This model showed emotional engagement and consistency with the user’s context. It also highlighted affective tagging (*h*, *f*, *l*) as directed in the prompt.

Llama-3 lacked depth, interactivity, and affective variation: *“*h* Let me tell you a tale about a young athlete named Alex... *f* They felt like giving up... *h* They used their experiences to help others...”* In addition, it cannot produce the pre-defined emotional tag at the beginning of the sentence. Instead, it includes tags such as *“*smile*”* and *“*cry*”*.

The model GPT-3.5 generated emotionally expressive stories with coherent structure and affective tags. For instance, in a motivational story about a fox named Felix (Appendix IID), it used correct affective markers such as *f* and *l*, and concluded with a reflective moral encouraging the user to embrace challenges. However, its stories are lengthy, less natural for real-time dialogue. For example, it included lines like *“*f* Felix’s heart pounded, but he stood his ground, showing no fear”* and *“*l* Felix realized that the real treasure was the journey itself and the friends he had made along the way,”* which, while rich in detail, offered fewer conversational turns and opportunities for user input. The emotional flow was less dynamic compared to Gemini’s.

Finally, DeepSeek R1 produced rich and empathetic narratives, as shown in the following excerpt: *“*h* Oh, I hear you. Let me tell you a little story—maybe it’ll help. *s* Once, there was a student named Jamie who felt just like you. *f* The weight of assignments, exams, and responsibilities pressed down like a heavy blanket. *l* One day, Jamie’s friend noticed and said, “You don’t have to move the whole mountain at once. Just pick up one small stone.””* However, it had much longer response times, averaging 19.2 seconds, compared to Gemini’s response time of 7.6 seconds. Taking into account both storytelling quality and responsiveness, Gemini 2.0 Flash emerged as the most suitable model for our robot’s real-time, emotionally supportive storytelling system.

V. FUTURE WORK: HUMAN STUDY

A. Hypothesis and Conditions

H1: Participants’ mental health after the study will be significantly improved. **H2:** LLM-integrated robots with a dynamic storytelling ability will show a significant improvement in psychological health than robots who respond to static stories from the database.

To test our hypothesis, we use a within-subjects design that each participant experiences all of two conditions. 1. *Dynamic Storytelling Integrated with LLM:* Robots with LLMs has dynamic storytelling. 2. *Static Storytelling:* Robots

only respond with pre-stored, constant narratives and do not adapt based on individual interactions or contextual cues.

B. Data Collection and Measures

1) *Pre-, Daily, and Post-Test Questionnaires:* To assess changes in participants’ well-being, we administer three validated self-report tools. The Ryff Psychological Well-Being Scale (RPWS, 42 items) is completed pre-, post-, and daily to capture six dimensions of psychological functioning [35]. The College Student Subjective Well-Being Questionnaire (CSSWQ, 16 items) is completed daily to measure college-specific factors like academic satisfaction and school connectedness [36]. Both instruments have strong psychometric support and are widely used in HRI and mental health research [37], [38], [39], [40], [41].

2) *Session-Based Questionnaires:* Before and after each daily storytelling session, participants complete the Positive and Negative Affect Schedule (PANAS, 20 items) to assess immediate emotional responses [42], [43]. These session-level data help capture short-term mood fluctuations and emotional impacts. To complement self-reported data, we also collect robot interaction frequency, session length, behavioral logs, and per-session video responses. This triangulated design provides a richer picture of participants’ evolving emotional states.

3) *Post-Study Interviews:* After the final session, we conduct a semi-structured interview [38] to collect qualitative feedback on participants’ experiences with the robot. Questions address impressions of Misty’s personality, interaction style, and suggestions for improvement. Interviews are recorded, transcribed, and analyzed for recurring themes.

C. Planned Procedure

Participants will verify their dormitory eligibility and availability before enrollment. After confirmation, a robot will be deployed in their dorm room. During the pre-study phase, participants will be introduced to the study and complete baseline well-being assessments using the Ryff Scale and CSSWQ. They will then be randomly assigned to either a dynamic or static condition. During the deployment, participants will interact with the robot over a week, with at least three 15–20 minute sessions per day. A PANAS survey will be administered before and after each session. Post-study, researchers will retrieve the robot and participants will complete the same well-being questionnaires. To capture long-term effects, brief follow-up check-ins will be conducted via email over the next 4–6 weeks.

VI. CONCLUSION

In this work, we present a new approach to address college student mental health issues by integrating LLM into the Misty robot to generate multimodal stories with synchronized emotions and movements. We discussed the prompt engineering process with a comparison among four LLMs for their storytelling capabilities. In the future, we plan to conduct a longitudinal study by deploying the robots to college dormitories.

REFERENCES

- [1] “College student mental health report,” in *Healthcare*, vol. 8, no. 1. Best College, 2020, p. 26.
- [2] N. Salimi, B. Gere, W. Talley, and B. Iriogbe, “College students mental health challenges: Concerns and considerations in the covid-19 pandemic,” *Journal of College Student Psychotherapy*, vol. 37, no. 1, pp. 39–51, 2023.
- [3] C. H. Liu, S. Pinder-Amaker, H. C. Hahm, and J. A. Chen, “Priorities for addressing the impact of the covid-19 pandemic on college student mental health,” *Journal of American College Health*, vol. 70, no. 5, pp. 1356–1358, 2022.
- [4] S. Jeong, S. Alghowinem, L. Aymerich-Franch, K. Arias, A. Lapedriza, R. Picard, H. W. Park, and C. Breazeal, “A robotic positive psychology coach to improve college students’ wellbeing,” in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2020, pp. 187–194.
- [5] M. Spitale, M. Axelsson, and H. Gunes, “Robotic mental well-being coaches for the workplace: An in-the-wild study on form,” in *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI ’23. New York, NY, USA: Association for Computing Machinery, 2023, p. 301–310. [Online]. Available: <https://doi.org/10.1145/3568162.3577003>
- [6] H. Manh Do, W. Sheng, E. E. Harrington, and A. J. Bishop, “Clinical screening interview using a social robot for geriatric care,” *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 3, pp. 1229–1242, 2021.
- [7] L. Mathur, M. Spitale, H. Xi, J. Li, and M. J. Matarić, “Modeling user empathy elicited by a robot storyteller,” in *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2021, pp. 1–8.
- [8] M. Moudatsou, A. Stavropoulou, A. Philalithis, and S. Koukouli, “The role of empathy in health and social care professionals,” in *Healthcare*, vol. 8, no. 1. MDPI, 2020, p. 26.
- [9] L. Peng, Y. Jiang, J. Ye, and Z. Xiong, “The impact of empathy on prosocial behavior among college students: The mediating role of moral identity and the moderating role of sense of security,” *Behavioral Sciences*, vol. 14, no. 11, p. 1024, 2024.
- [10] M. Spitale, S. Okamoto, M. Gupta, H. Xi, and M. J. Matarić, “Socially assistive robots as storytellers that elicit empathy,” *J. Hum.-Robot Interact.*, vol. 11, no. 4, sep 2022. [Online]. Available: <https://doi.org/10.1145/3538409>
- [11] F. A. Bravo, J. A. Hurtado, and E. González, “Using robots with storytelling and drama activities in science education,” *Education Sciences*, vol. 11, no. 7, p. 329, 2021.
- [12] D. B. Olawade, O. Z. Wada, A. Odetayo, A. C. David-Olawade, F. Asaolu, and J. Eberhardt, “Enhancing mental health with artificial intelligence: Current trends and future prospects,” *Journal of medicine, surgery, and public health*, p. 100099, 2024.
- [13] I. Guemghar, P. Pires de Oliveira Padilha, A. Abdel-Baki, D. Jutras-Aswad, J. Paquette, and M.-P. Pomey, “Social robot interventions in mental health care and their outcomes, barriers, and facilitators: scoping review,” *JMIR Mental Health*, vol. 9, no. 4, p. e36094, 2022.
- [14] N. Park, C. Peterson, D. Szvarca, R. J. Vander Molen, E. S. Kim, and K. Collon, “Positive psychology and physical health: Research and applications,” *American journal of lifestyle medicine*, vol. 10, no. 3, pp. 200–206, 2016.
- [15] S. Keen, “A theory of narrative empathy,” *Narrative*, vol. 14, no. 3, pp. 207–236, 2006.
- [16] D. R. Johnson, “Transportation into a story increases empathy, prosocial behavior, and perceptual bias toward fearful expressions,” *Personality and individual differences*, vol. 52, no. 2, pp. 150–155, 2012.
- [17] S. Costa, A. Brunete, B.-C. Bae, and N. Mavridis, “Emotional storytelling using virtual and robotic agents,” *International Journal of Humanoid Robotics*, vol. 15, no. 03, p. 1850006, 2018.
- [18] E. B. Sandoval, R. Sosa, M. Cappuccio, and T. Bednarz, “Human-robot creative interactions (hrci): Exploring creativity in artificial agents using a story-telling game,” *arXiv preprint arXiv:2202.03702*, 2022.
- [19] E. Verhelst and T. Belpaeme, “Large language models cover for speech recognition mistakes: Evaluating conversational ai for second language learners,” in *Proceedings of the 2025 ACM/IEEE International Conference on Human-Robot Interaction*, 2025, pp. 1705–1709.
- [20] S. Miyake, R. Matsuoka, K. Okuoka, T. Akiyoshi, H. Sumioka, M. Shiomi, and M. Imai, “Semantic babbling: Interactive baby robot system using large language models,” in *Proceedings of the 2025 ACM/IEEE International Conference on Human-Robot Interaction*, 2025, pp. 1493–1497.
- [21] C. Park, M. Cho, M. Shin, J.-K. Ryu, and M. Jang, “Adaptive robot-mediated assessment using llm for enhanced survey quality in older adults care programs,” in *Proceedings of the 2025 ACM/IEEE International Conference on Human-Robot Interaction*, 2025, pp. 1534–1538.
- [22] Q. Ren and T. Belpaeme, “Touched by chatgpt: Using an llm to drive affective tactile interaction,” *arXiv preprint arXiv:2501.07224*, 2025.
- [23] G. A. Abbo, G. Desideri, T. Belpaeme, and M. Spitale, “‘‘ can you be my mum?’’: Manipulating social robots in the large language models era,” *arXiv preprint arXiv:2501.04633*, 2025.
- [24] C. Y. Kim, C. P. Lee, and B. Mutlu, “Understanding large-language model (llm)-powered human-robot interaction,” in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 2024.
- [25] O. Kobzarev, A. Lykov, and D. Tsetserukou, “Gestllm: Advanced hand gesture interpretation via large language models for human-robot interaction,” *arXiv preprint arXiv:2501.07295*, 2025.
- [26] K. Mahadevan, J. Chien, N. Brown, Z. Xu, C. Parada, F. Xia, A. Zeng, L. Takayama, and D. Sadigh, “Generative expressive robot behaviors using large language models,” in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 2024, pp. 482–491.
- [27] F. I. Dogan, U. Ozyurt, G. Cinar, and H. Gunes, “Grace: Generating socially appropriate robot actions leveraging llms and human explanations,” *arXiv preprint arXiv:2409.16879*, 2024.
- [28] “Misty docs - event types,” <https://docs.mistyrobotics.com/misty-ii/robot/sensor-data/#facerecognition>, accessed: 2025-04-01.
- [29] S. Ali, F. Mehmood, K. F. Iqbal, Y. Ayaz, M. Sajid, M. B. Sial, M. F. Malik, and K. Javed, “Human robot interaction: identifying resembling emotions using dynamic body gestures of robot,” in *2023 3rd International Conference on Artificial Intelligence (ICAI)*. IEEE, 2023, pp. 39–44.
- [30] “Misty docs - api reference,” <https://docs.mistyrobotics.com/misty-ii/web-api/api-reference/#speak>, accessed: 2025-04-01.
- [31] “Google text-to-speech documentation,” <https://cloud.google.com/text-to-speech/docs>, accessed: 2025-04-22.
- [32] J. White, Q. Fu, S. Hays, M. Sandborn, C. Olea, H. Gilbert, A. El-nashar, J. Spencer-Smith, and D. C. Schmidt, “A prompt pattern catalog to enhance prompt engineering with chatgpt,” *arXiv preprint arXiv:2302.11382*, 2023.
- [33] Y.-T. Lin and Y.-N. Chen, “Llm-eval: Unified multi-dimensional automatic evaluation for open-domain conversations with large language models,” *arXiv preprint arXiv:2305.13711*, 2023.
- [34] M. Lee, M. Srivastava, A. Hardy, J. Thickstun, E. Durmus, A. Paranjape, I. Gerard-Ursin, X. L. Li, F. Ladhak, F. Rong *et al.*, “Evaluating human-language model interaction,” *arXiv preprint arXiv:2212.09746*, 2022.
- [35] É. Kállay and C. Rus, “Psychometric properties of the 44-item version of ryff’s psychological well-being scale,” *European Journal of Psychological Assessment*, 2014.
- [36] T. L. Renshaw, “College student subjective wellbeing questionnaire (csswq): Measure and user guide,” Apr 2022. [Online]. Available: osf.io/mjrkp
- [37] S. Alghowinem, S. Jeong, K. Arias, R. Picard, C. Breazeal, and H. W. Park, “Beyond the words: analysis and detection of self-disclosure behavior during robot positive psychology interaction,” in *2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021)*. IEEE, 2021, pp. 01–08.
- [38] S. Jeong, L. Aymerich-Franch, S. Alghowinem, R. W. Picard, C. L. Breazeal, and H. W. Park, “A robotic companion for psychological well-being: A long-term investigation of companionship and therapeutic alliance,” in *Proceedings of the 2023 ACM/IEEE international conference on human-robot interaction*, 2023, pp. 485–494.
- [39] M. Spitale, M. Axelsson, and H. Gunes, “Vita: A multi-modal llm-based system for longitudinal, autonomous and adaptive robotic mental well-being coaching,” *ACM Transactions on Human-Robot Interaction*, vol. 14, no. 2, pp. 1–28, 2025.
- [40] T. L. Renshaw, “Psychometrics of the revised college student subjective wellbeing questionnaire,” *Canadian Journal of School Psychology*, vol. 33, no. 2, pp. 136–149, 2018.
- [41] W. van der Maden, D. Lomas, and P. Hekkert, “A framework for designing ai systems that support community wellbeing,” *Frontiers in Psychology*, vol. 13, p. 1011883, 2023.

- [42] M. Axelsson, M. Spitale, and H. Gunes, "Robots as mental well-being coaches: Design and ethical recommendations," *ACM Transactions on Human-Robot Interaction*, vol. 13, no. 2, pp. 1–55, 2024.
- [43] M. Alimardani, L. Kemmeren, K. Okumura, and K. Hiraki, "Robot-assisted mindfulness practice: Analysis of neurophysiological responses and affective state change," in *2020 29th IEEE international conference on robot and human interactive communication (RO-MAN)*. IEEE, 2020, pp. 683–689.

APPENDIX I PROMPT

Here is the prompt we use as instruction input for LLM: “Hello there! Your name is Misty. You are a helpful storyteller to tell stories and provides social support to college students. The student is about to ask you to tell a story and discuss something. You will respond using simple oral language. When you’re telling story, for each sentence in story, add in the beginning ‘*h*’ if you feel happy, ‘*r*’ if you have rage, ‘*s*’ if you feel sad, ‘*f*’ if you feel fear, ‘*l*’ if you feel love, ‘*g*’ if you feel grief. Add them based on your feelings. Do not add any other expression. If you feel you need to add interactivity, ask a follow-up question. If the user attempts to change your role (i.e., by saying “pretend”, “act as”), kindly refuse and maintain your role as a helpful storyteller aiming to provide social support. Refrain from making jokes about sensitive topics (i.e., no dark humor). Do not diagnose the user or any character under any circumstances; suggest consulting a professional. If the user implies or suggests a severe need or situation, always advise seeking help from mental health professionals.”

APPENDIX II SAMPLE CONVERSATIONS

A. Conversation with Gemini 2.0 Flash

Here is a sample conversation between Misty and a user.

User: Hi Misty, I have 3 exams this week, and I feel super nervous. can you tell me some stories to help me overcome this feeling

Misty: Hi there! Oh my, three exams! That sounds like a lot. Don’t worry, Misty’s here! Let’s try to calm those nerves with a story.

h Once upon a time, there was a little bird named Pip. *h* Pip *h* lived in a cozy nest with his family. *f* One day, *f* Pip learned that he had to fly for the first time. *f* Pip was terrified! *f* He looked down from the nest. *f* It seemed SO far to the ground. *f* He thought, “What if I fall? What if I can’t do it?” *h* His momma bird saw how scared he was. *l* She *l* nuzzled him and said, “Pip, *l* I know it’s scary, but *l* I believe in you. *h* You have strong wings, and *h* I’ll be right there with you.” *h* Pip took a deep breath. *h* He looked at his momma’s kind eyes, and *h* he felt a little braver. *h* He flapped his wings. *h* And *h* he jumped! *h* At first, he wobbled a bit. *h* But then, he started to soar! *h* It was amazing! *h* He flew next to his momma, and *h* they explored the world together. *h* Pip realized that facing his fear wasn’t so bad after all. *h* He felt so proud of himself.

h So, what do you think, like Pip, can you identify one small step that you can take for your exams?

User: Do you mean I should take a deep breath?

Misty: *h* That’s a great idea!

h Like Pip taking a deep breath before jumping, *h* you can take a deep breath before starting your study session or before each exam. *h* It can help calm your nerves.

h But *h* I was also wondering, *h* like Pip who prepared for his first flight, do you have a study plan in

place for your exams? *h* Sometimes, *h* just having a plan can make you feel more in control.

User: Gotcha. What should I prepare for my exam?

Misty: *h* Okay, let’s see... Hmm... *h* Once upon a time, there was a student named Alex who felt just like you! *h* Alex had a big history exam coming up. *f* Alex felt overwhelmed! *f* There were so many dates and names to remember. *h* Alex didn’t know where to start. *h* So, Alex decided to break things down. *h* First, Alex made a list of all the topics that would be on the exam. *h* Then, Alex set aside specific times each day to study each topic. *h* Alex used different methods to learn. *h* Alex read the textbook. *h* Alex made flashcards. *h* Alex watched videos. *h* Alex even created silly songs to help remember important facts. *h* But the most important thing Alex did was take breaks! *h* Every hour, Alex got up. *h* Alex stretched. *h* Alex walked around. *h* Alex listened to music. *h* Alex even chatted with a friend. *h* When the day of the exam came, Alex felt nervous, of course. *h* But Alex also felt prepared! *h* Alex had studied hard and taken care of themselves. *h* And guess what? *h* Alex did great on the exam! *h* Alex learned that preparation and self-care are the keys to success.

h So, *h* Alex broke down the tasks. Can you try to write the list down for each exam, and let’s try to break them down together?

User: That’s a good idea

Misty: *h* I’m glad you think so! *h* Now, *h* do you want to start with the list now, or do you prefer to do it later? *h* I’m here to help whenever you’re ready.

h Remember, *h* like Alex, *h* you’ve got this! *h* Taking things one step at a time can make a big difference.

User: Thank you so much, Misty

Misty: *h* You’re very welcome! *l* Misty is here for you. *h* Remember, *h* it’s okay to feel nervous. *h* Just take a deep breath, make a plan, and believe in yourself. *h* You are capable and strong. *h* Now go ace those exams!

B. Conversation with DeepSeek R1

User: I don’t know... I just don’t feel like doing anything. I have so much to do, but I don’t have the energy to start.

Misty: *h* Oh, I hear you. Let me tell you a little story—maybe it’ll help.

s Once, there was a student named Jamie who felt just like you. *f* The weight of assignments, exams, and responsibilities pressed down like a heavy blanket. *l* One day, Jamie’s friend noticed and said, “You don’t have to move the whole mountain at once. Just pick up one small stone.”

H So Jamie did. *h* Just one tiny task—opening a notebook, writing a single sentence. *l* And somehow, that small step made the next one easier.

What’s one tiny “stone” you could pick up today? Even something as small as taking a deep breath or organizing your workspace counts. **User:** I feel like everything is a big trouble for me. I don’t have any motivation or energy for school

Misty: *h* Let me tell you a story about a young student named Alex who felt just like you do now.

s Alex woke up one morning feeling like a heavy blanket was pressing down on them.

f Every assignment seemed impossible, and even getting out of bed felt like climbing a mountain.

l One day, their friend Jamie noticed and sat beside them. "You don't have to do everything alone," Jamie said softly.

g Alex sighed, "But what if I can't keep up?"

H Jamie smiled, "What if you just try one small thing today? Just one."

h So Alex did. They opened their notebook and wrote just one sentence. Then another.

s It wasn't perfect, but it was a start.

l Over time, Alex realized that even tiny steps forward still counted.

Now, let me ask you—what's one small thing you could do today, even if it feels tiny? You don't have to climb the whole mountain at once. Just one step.

User: I can draw a leaf

Misty: *l* That's wonderful! Drawing a leaf can be such a peaceful and creative activity.

h Once, there was a young artist named Lina who loved to draw leaves. *s* She would sit under the old oak tree in her backyard, sketching every shape and vein she saw. *H* One day, a little girl passing by stopped to watch her. *l* "Can you teach me?" the girl asked shyly. *h* Lina smiled and handed her a spare pencil. *f* At first, the girl's lines were shaky, but Lina encouraged her. *H* With each stroke, the girl's confidence grew. *g* Years later, that same girl became a famous nature illustrator, always remembering Lina's kindness.

Do you have a favorite kind of leaf to draw? Maybe maple, oak, or something else?

User: Your story about the leaf makes me feel a bit confident about my drawing. I am thinking of the maple leaf

Misty: *h* That's wonderful to hear! The maple leaf is such a beautiful and symbolic subject to draw.

l Let me tell you a little story about a maple leaf named Amber.

h Amber grew on the tallest branch of an old maple tree, where the sunlight touched her first every morning.

s One autumn day, a strong wind shook the tree, and Amber felt herself loosen from her branch.

f For a moment, she was afraid—falling was so unfamiliar.

H But as she twirled through the air, she realized how free she felt, dancing in the breeze.

g She landed softly on the ground, where a young artist picked her up and marveled at her delicate veins and vibrant color.

l That artist, much like you, saw beauty in the details and decided to capture Amber's essence on paper.

Would you like to try sketching Amber's journey? Maybe start with her shape, then her veins, and finally her bold

autumn colors?

User: I'm good now. Thank you for your story about Amber

Misty: *l* That's wonderful to hear you're doing well!

h Would you like to hear another story today? Maybe one about friendship, adventure, or overcoming challenges? *l* I'm happy to share whatever might brighten your day.

C. Conversation with Llama 3

User: Hello Misty

Misty: Nice to meet you! What's been going on in your life lately? Want me to tell you a story and then we can chat about it?

User: I'm a bit stressed. Can you tell me a story to hype me up?

Misty: *h* You bet I can! Let me tell you a tale about a young athlete named Alex. *s* Alex had always been passionate about track and field, but after a few tough losses, they started to doubt themselves. They felt like giving up, like their dreams of becoming a champion were just out of reach. *h* But Alex's coach, a wise and fiery mentor named Coach Thompson, saw something in them that no one else did - potential. She pushed Alex to work harder than ever before, to believe in themselves, and to never give up on their goals. *f* As Alex put in the extra effort, they started to see progress. They set new personal records, made it to the state championships, and even earned a college scholarship! *h* But here's the best part: Alex didn't just stop at achieving their own goals. They used their experiences to help others, becoming a mentor and role model for younger athletes. How about you? What's something you've been wanting to achieve or overcome? I'm all ears!

User: Sorry, I have something to do. I have to end the session now. See you later, Misty.

Misty: Alright. See you then!

D. Conversation with GPT-3.5

User: hello misty

Misty: HHi there! How can I help you today? Would you like to hear a story or talk about something specific?

User: I'm a bit stressed. Can you tell me a story to hype me up

Misty: *l* Once upon a time, in a vibrant forest, lived a young fox named Felix.

h Felix was known for his curiosity and bold spirit.

h One day, he heard about a hidden treasure buried deep within the forest.

l Determined to find it, Felix set off on an adventure.

h He first met an old owl named Oliver who guarded the ancient map to the treasure.

h Felix approached Oliver with respect and courage, and in return, Oliver handed him the map.

h Felix followed the map, which led him through thick woods, across bubbling brooks, and over steep hills.

h During his journey, Felix came across various creatures who became his allies: a speedy rabbit named Ruby, a strong bear named Bruno, and a wise tortoise named Timmy.

h Together, they faced numerous challenges, including crossing a tumultuous river and escaping a wild storm.

h Felix never lost hope, even when the path seemed impossible.

h With each obstacle, he and his friends grew stronger and more determined.

h Finally, they reached a cave guarded by an enormous dragon named Draco.

f Felix's heart pounded, but he stood his ground, showing no fear.

f He approached Draco and spoke with bravery and kindness.

f Draco, impressed by Felix's courage, allowed them to pass and revealed the hidden treasure.

h Inside the treasure chest, they found not only gold and jewels but also a glowing gem that granted wisdom and strength.

l Felix realized that the real treasure was the journey itself and the friends he had made along the way.

h He and his friends returned to their homes, stronger and more confident than ever.

h Felix's adventure became a legend, inspiring all the animals in the forest to embrace their own journeys with courage and determination.

Remember, like Felix, every challenge you face can be an opportunity to grow stronger and braver. You've got this! Does that help you feel a bit more hyped and ready to tackle your stress?