Exploring the Utilities of LLM's in Robot-Supported Mindfulness Practices

SHRIRANG PATIL, Oregon State University, USA
THEING MWE OO, Oregon State University, USA
HEATHER KNIGHT, Oregon State University, USA

Fig. 1. Generated image depicting a Pepper-Inspired Robot Meditating with People in Indoor/Outdoor Space [Dall-E w/ Seph Divine]

1 ABSTRACT

In mindfulness practices, many struggle to adhere to regular routines. Socially assistive robots represent promising guiding meditation sessions to promote engagement. This research investigates integrating responsiveness powered by recent large language models (LLM) into robot-led mindfulness. We present two LLM-enhanced robot modalities: first, an LLM-driven pipeline generating tailored verbal meditation guidance performed by a Pepper robot; second, a pipeline detecting human emotions from speech, them mapping them to robot gestural response. Initial implementations indicate a promising subject response. A six-session robot-facilitated meditation series occurring on consecutive new moons will inform ongoing innovation, contributing to a rising domain of automated robot assistants for well-being practices. Later work will use LLM analysis of human speech, e.g., mapping affective content to appropriate robot gestural responses. We invite others to harness modern LLMs' potential for human-robot storytelling, promoting a caring, playful side of this technology, and supporting groups of humans in joint practice for this important skill.

Additional Key Words and Phrases: Social robots, robot meditation, technology-supported mindfulness, affective speech analysis

ACM Reference Format:

Shrirang Patil, Theing Mwe Oo, and Heather Knight. 2024. Exploring the Utilities of LLM's in Robot-Supported Mindfulness Practices. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24 Companion), March 11–14, 2024, Boulder, CO, USA*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3610978.3640607

Manuscript submitted to ACM

2 INTRODUCTION

Mindfulness practices, such as meditation and yoga, have demonstrated mental and physical health benefits. However, the challenge lies in starting and adhering to regular routines. Socially assistive robots show promise as coaches guiding audiences through sessions to promote accessibility and motivation. For instance, Axelsson et al. [3] successfully deployed robots in a public cafe to lead group breathing awareness practices.

Recent advances in large language models (LLMs) have also opened up new possibilities for enhancing robot-supported mindfulness. Kumar et al. [13] found that LLM chatbots improved people's intent to practice mindfulness, highlighting the conversational agents' role. Additionally, Spitale et al. [21] compared physical and virtual robot speech therapy coaches, finding the physical robot to deliver better child engagement and linguistic gains.

This research aims to investigate utilities of large language models to enhance robot coaches for mindfulness, via the exploration of two specific LLM robot modalities: (1) A robot that uses LLMs to generate guided meditation scripts to be delivered out-loud to a human audience, (2) A robot that performs affective human speech recognition to offer comforting gestures via sentiment analysis. These implementations will help us explore our research question: *What can LLMs offer to HRI-related Mindfulness Practices?*

3 RELATED WORK

Although the use of robots and speechbots in mindfulness therapy are not yet common, the first late breaking reports [3, 8, 15, 20, 21] and full papers [2, 4, 7] show promising first steps. For example, Axelsson et al. [3] found café customers willing to participate in robot-led breathing awareness sessions. Bodala et al. [4] compared how users perceived mindfulness training delivered by a robot versus a human. [15] developed a robot-guided deep breathing practice to reduce anxiety. [2, 12, 22] studies have found positive perceptions, experience impacts, and neurophysiological changes from robot-guided mindfulness sessions that are comparable to human-delivered practices.

Broader uses of robots in therapeutic and human-supporting applications range from health and exercise support [5, 6, 14] to physical [21] and emotional therapy [11, 12, 16]. These adjacent domains sometimes use metrics like therapeutic alliance, participant desire to continue doing exercises with the robot, challenges of this varieties, and self-reported health outcomes. For instance, Matheus et al. [14] detected breathing phases for personalized interactions while Spitale et al. [21] compared child engagement between physical and virtual robot speech therapy assistants. [11] assessed psychological well-being improvements in students interacting with a supportive robot coach. Moreover, Robinson et al. [17] found that higher education students had a moderately good perception of a humanoid robot delivering a mindful breathing technique. So much interest, but few with repeated sessions or community.

Algorithmic therapy systems are also on the rise. Denecke et al. [7] presented SERMO, a mental health chatbot automatically suggesting coping strategies based on detected emotions. Similarly, [9] EMMA- The ChatBot was designed to detect mood from phone sensors and provide relevant wellbeing micro-interventions. As language capabilities advance, there will be growing potential to integrate features like natural language processing, emotion recognition, and personalization[9, 18] into socially assistive robots to optimize their coaching interactions and outcomes.

4 PROBLEM STATEMENT

To investigate the potential of LLMs in supporting human mindfulness practices, we propose to develop and evaluate two human-robot mindfulness applications. In the first, we will have a robot / voicebot lead spoken meditation (the first series of which occurs on New Moon). In the second, we propose a compassionate listener robot that will generate

empathetic gestures; for instance, calming motions would be generated if elevated stress or anger is detected through human speech, aiming to promote relaxation and well-being through nonverbal signals.

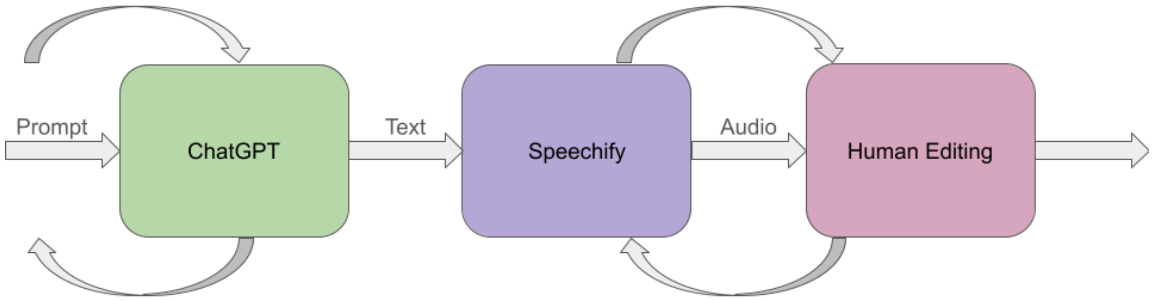

Fig. 2. The LLM-Speechify robot guided meditation pipeline

Fig (2) depicts a pipeline for robot-guided mindfulness using ChatGPT text converted into robotic voices with the help of the Speechify AI tool and human-in-the-loop editing. As a final step, calm music was added in the background. In addition to speech, the Pepper robot performed programmed motions. Systems such as this one might be used to guide practitioners through personalized mindfulness interventions to promote engagement and motivation in future work, but for now, we are using them in a community event.

Fig. 3. The human speech emotion recognition to robot gesture mapping pipeline

Fig (3) depicts a concept design for processing human speech into emotion, and then generating empathetic robot gestures. This research would explore the integration of LLM capabilities to enhance a robot's responsive verbal guidance and non-verbal cues. The motivation is to create robots capable of more compassionate, supportive behaviors to nurture better bonds and outcomes in human-robot interaction. This relates closely to ongoing research in emotionally intelligent, socially assistive robotics - an emerging domain with application in areas like mindfulness, yoga and mental health and more.

5 FIRST ROBOT NEW MOON MEDIATION

In the kickoff to our monthly robot guided mindfulness sessions, we leveraged LLMs to generate a 4.5 minute mindfulness coaching dialogue to be delivered by a Pepper robot (Fig. 2). The theme of this first edition was "*Robot Harmony*" and it was attended by four graduate students, a professor and a robot. Two voicebots (one male and one female) were selected by the second author to deliver segments of the script. Prompt engineering was used to prime ChatGPT to include robots, the solar system, and new moon intention setting into a 500-word guided meditation script that was human edited for accuracy (e.g., new moons do not emit light) and comedic flow. Here is a segment:

> *Let the gentle whirring of gears become your guide, leading you deeper into the present moment. Now, turn your attention to the night sky, envisioning the moon as a powerful motherboard radiating cosmic energy. The new moon, a symbol of renewal, offers us a blank canvas upon which to project our aspirations.*
>
> *Just as robots are programmed for specific tasks, this lunar phase empowers us to program our intentions into the universe. Feel the energy of the new moon washing over you, cleansing your spirit and preparing you for the journey ahead...Imagine the robots around you absorbing this lunar energy, transforming it into a harmonious force that resonates within and around you.*

The group spent about 45 minutes together, opening and closing with the generative voice and motion script, in which both the spoken meditation and Pepper robot motions were cued at the same time. The group was encouraged to set intentions that they could watch grow with the moon over the next two weeks and instructed to observe the moon over its 30 day cycle, also considering what to let go of as the full moon reduces to a crescent. The group shared their first impressions of the system and experience, and some shared their plans for the upcoming 30-day cycle.

6 TOWARDS EMPATHETIC ROBOT LISTENERS

In our upcoming technology development, we plan to design a robot that can respond to human speech with empathetic gestures, investigating both affect analysis and context-to-gesture mappings. One thought is to conduct word or prosody analysis to classify if there are strong emotions present, e.g., anxiety, frustration or calmness. Next, we will research, develop, and evaluate relevant robot gestures to respond to these emotion types, perhaps augmented by the human animator's feedback.

For example, detecting agitation may trigger calming arm motions and downcast eyes. Detecting interest could produce affirming postures and open robot body language (e.g., wide arms with open palms). Findings could inspire more ambitious efforts to grow machine compassion through fusing speech understanding and physical actions.

7 DISCUSSION/CONCLUSION

In the context of emotional vulnerability, LLM-based social robots introduce a multifaceted array of potential risks due to their statistical nature and lack of genuine understanding. These models, while capable of generating unpredictable responses without a consistent emotional foundation, pose a risk of inappropriate or potentially harmful interactions. Despite these challenges, research, including work by Spitale et al.[19], highlights a potential benefit in encouraging individuals to cultivate a relationship with themselves, potentially fostering self-reflection and personal growth through practices such as well-being coaching and mindfulness. This nuanced perspective underscores the importance of carefully navigating the integration of LLMs in emotionally sensitive applications, balancing the innovative potential against the ethical and practical risks involved.

Thus far, we have utilized ChatGPT to generate a tailored mindfulness coaching text for delivery through a Pepper robot's voice. The script guided the audience using Speechify's text-to-speech. In our ongoing work, we would like to match speech to robot motion in both application cases, expanding repertoires of mindfulness activities. It is common at meditation retreats to engage in sharing circles, or meet 1:1 with guides alongside sitting, walking, or other more well known mindfulness practices. Having opportunities in which humans can share, wherein a robot guide had suitable nonverbal responses may increase people's reflections about and commitment to mindfulness.

We would also like to explore using LLM-based speech emotion recognition to detect the user's engagement and reactions during the session (as in Fig. 3), responding to human utterances with empathetic gestures. Generating suitable

robot gestures and motions have also been shown to be helpful in other robot coaching domains: Gjaci et al.[10] explores learning culturally-specific co-speech gestures from audio-pose mappings, showcasing potential to capture nuances. Complementary to this, Alexanderson et al.[1] develops controllable diffusion models for dance motion synthesis from audio. For instance, signs of interest and relaxation may one day trigger encouraging robot gestures. Our goal is to use LLM mindfulness to help people engage with their own inner workings.

ACKNOWLEDGMENTS

To Luke Sanchez who crafted the robot motion, and to the LLM cosmic robots whose gears have a soothing hum.

REFERENCES

[1] Simon Alexanderson, Rajmund Nagy, Jonas Beskow, and Gustav Eje Henter. 2022. Listen, Denoise, Action! Audio-Driven Motion Synthesis with Diffusion Models. *ACM Transactions on Graphics (TOG)* 42 (2022), 1 – 20. https://api.semanticscholar.org/CorpusID:253581728

[2] Maryam Alimardani, Linda Kemmeren, Kazuki Okumura, and Kazuo Hiraki. 2020. Robot-Assisted Mindfulness Practice: Analysis of Neurophysiological Responses and Affective State Change. *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)* (2020), 683–689. https://api.semanticscholar.org/CorpusID:221104010

[3] Minja Axelsson, Micol Spitale, and Hatice Gunes. 2023. Robotic Coaches Delivering Group Mindfulness Practice at a Public Cafe. *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction* (2023). https://api.semanticscholar.org/CorpusID:257406411

[4] Indu Prasad Bodala, Nikhil Churamani, and Hatice Gunes. 2021. Teleoperated Robot Coaching for Mindfulness Training: A Longitudinal Study. *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)* (2021), 939–944. https://api.semanticscholar.org/CorpusID:237297069

[5] Indu Prasad Bodala and Hatice Gunes. 2021. Dynamic Bayesian Network Modelling of User Affect and Perceptions of a Teleoperated Robot Coach during Longitudinal Mindfulness Training. *ArXiv* abs/2112.02017 (2021). https://api.semanticscholar.org/CorpusID:244896131

[6] Jiaee Cheong, Micol Spitale, and Hatice Gunes. 2023. "It's not Fair!" – Fairness for a Small Dataset of Multi-modal Dyadic Mental Well-being Coaching. *2023 11th International Conference on Affective Computing and Intelligent Interaction (ACII)* (2023), 1–8. https://api.semanticscholar.org/CorpusID:263677413

[7] Kerstin Denecke, Sayan Vaaheesan, and Aaganya Arulnathan. 2020. A Mental Health Chatbot for Regulating Emotions (SERMO) - Concept and Usability Test. *IEEE Transactions on Emerging Topics in Computing* 9 (2020), 1170–1182. https://api.semanticscholar.org/CorpusID:213810982

[8] Ning Fang, Chao Zhang, Supraja Sankaran, and Shaoya Ren. 2022. Role of Socially Assistive Robots in Reducing Anxiety and Preserving Autonomy in Children. *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2022), 754–759. https://api.semanticscholar.org/CorpusID:247619375

[9] Asma Ghandeharioun, Daniel J. McDuff, Mary Czerwinski, and Kael Rowan. 2018. EMMA: An Emotion-Aware Wellbeing Chatbot. *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)* (2018), 1–7. https://api.semanticscholar.org/CorpusID:198179485

[10] Ariel Gjaci, Carmine Tommaso Recchiuto, and Antonio Sgorbissa. 2022. Towards Culture-Aware Co-Speech Gestures for Social Robots. *International Journal of Social Robotics* 14 (2022), 1493 – 1506. https://api.semanticscholar.org/CorpusID:249353761

[11] Sooyeon Jeong, Sharifa Alghowinem, Laura Aymerich-Franch, Kika Arias, Àgata Lapedriza, Rosalind Picard, Hae Won Park, and Cynthia Lynn Breazeal. 2020. A Robotic Positive Psychology Coach to Improve College Students' Wellbeing. *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)* (2020), 187–194. https://api.semanticscholar.org/CorpusID:221534231

[12] Sooyeon Jeong, Laura Aymerich-Franch, Sharifa Alghowinem, Rosalind W. Picard, Cynthia Lynn Breazeal, and Hae Won Park. 2023. A Robotic Companion for Psychological Well-being: A Long-term Investigation of Companionship and Therapeutic Alliance. *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction* (2023). https://api.semanticscholar.org/CorpusID:257430665

[13] Harsh Kumar, Yiyi Wang, Jiakai Shi, Ilya Musabirov, Norman A. S. Farb, and Joseph Jay Williams. 2023. Exploring the Use of Large Language Models for Improving the Awareness of Mindfulness. *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (2023). https://api.semanticscholar.org/CorpusID:258217807

[14] Kayla Matheus, Ellie Mamantov, Marynel Vázquez, and Brian Scassellati. 2023. Deep Breathing Phase Classification with a Social Robot for Mental Health. *Proceedings of the 25th International Conference on Multimodal Interaction* (2023). https://api.semanticscholar.org/CorpusID:263742971

[15] Kayla Matheus, Marynel Vázquez, and Brian Scassellati. 2022. A Social Robot for Anxiety Reduction via Deep Breathing. *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)* (2022), 89–94. https://api.semanticscholar.org/CorpusID:251673077

[16] Aurea Bravo Perucho and Maryam Alimardani. 2023. Social Robots in Secondary Education: Can Robots Assist Young Adult Learners with Math Learning? *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction* (2023). https://api.semanticscholar.org/CorpusID:257406249

[17] Nicole L. Robinson, Jennifer F. Connolly, Gavin Suddrey, and David John Kavanagh. 2023. A Brief Wellbeing Training Session Delivered by a Humanoid Social Robot: A Pilot Randomized Controlled Trial. *International Journal of Social Robotics* (2023), 1–15. https://api.semanticscholar.org/CorpusID:

260887052

[18] Zhonghao Shi, Han Chen, Anna-Maria Velentza, Siqi Liu, Nathaniel Dennler, Allison O'Connell, and Maja J. Matarić. 2023. Evaluating and Personalizing User-Perceived Quality of Text-to-Speech Voices for Delivering Mindfulness Meditation with Different Physical Embodiments. *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction* (2023). https://api.semanticscholar.org/CorpusID:257430695

[19] Micol Spitale, Minja Axelsson, and Hatice Gunes. 2023. VITA: A Multi-modal LLM-based System for Longitudinal, Autonomous, and Adaptive Robotic Mental Well-being Coaching. *ArXiv* abs/2312.09740 (2023). https://api.semanticscholar.org/CorpusID:266335775

[20] Micol Spitale, Minja Axelsson, Neval Kara, and Hatice Gunes. 2023. Longitudinal Evolution of Coachees' Behavioural Responses to Interaction Ruptures in Robotic Positive Psychology Coaching. *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)* (2023), 315–322. https://api.semanticscholar.org/CorpusID:263162306

[21] Micol Spitale, Silvia Silleresi, Franca Garzotto, and Maja J. Matarić. 2023. Using Socially Assistive Robots in Speech-Language Therapy for Children with Language Impairments. *International Journal of Social Robotics* 15 (2023), 1525–1542. https://api.semanticscholar.org/CorpusID:260719646

[22] Sue Yoon, Maryam Alimardani, and Kazuo Hiraki. 2021. The Effect of Robot-Guided Meditation on Intra-Brain EEG Phase Synchronization. *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction* (2021). https://api.semanticscholar.org/CorpusID:232136490