Designing Culturally Adaptive Emotional Gestures to Enhance Child-Robot Interaction with NAO Robots in ASD Therapy

Chan Sri Manukalpa
Department of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
chansrimanukalpa@gmail.com

Koliya Pulasinghe
Department of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
koliya.p@sliit.lk

Samantha Rajapakshe
Department of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
samantha.r@sliit.lk

2025 International Research Conference on Smart Computing and Systems Engineering (SCSE) | 979-8-3315-2326-8/25/$31.00 ©2025 IEEE | DOI: 10.1109/SCSE65633.2025.11031067

Abstract—Integrating social robots into human-robot interactions demands advancements in natural language processing, navigation, computer vision, and expressive gestures to foster meaningful interactions. However, a gap remains in designing culturally relevant and developmentally appropriate gestures, particularly in the Sri Lankan context. Autism Spectrum Disorder (ASD), a neurodevelopmental condition impacting early education, often remains underdiagnosed, exacerbating learning challenges. This study introduces a novel approach utilizing robot-child interactions for ASD screening to minimize such delays. Expressive gestures were developed for the NAO6 humanoid robot to engage Sinhala-speaking children aged 2 to 6 years, including those with ASD, in Sri Lanka. Using the NAOqi Python API and Choregraphe simulator, culturally aligned gestures expressing emotions like happiness, sadness, fear, anger, and more were designed and synchronized with voice and LED effects. Pilot studies with typical children demonstrated the significance of linguistic and cultural alignment in enhancing engagement, emotional response, and trust. By addressing cultural nuances and advancing early ASD screening, this framework holds potential for broader applications in education, therapy, and diagnosis, improving human-robot interactions globally.

Keywords—autism spectrum disorder, Human-Robot Interaction, gesture synthesis

I. Introduction (*Heading 1*)

Social robotics encompasses various areas that facilitate the integration of robots into society and enhance their interactions with humans. This includes exploring Natural Language Processing for verbal communication with individuals and utilizing computer vision for personalized behaviors through perception, such as recognizing faces or gauging user engagement. Additionally, navigation capabilities are essential for robots to move independently within their environments. Gestures make interactions with social robots feel more natural and enjoyable. Research has shown that a robot's varied gestures and vocal tones can significantly influence users' emotional responses [1].

Autism significantly affects school-aged children in Sri Lanka, underscoring the need for culturally sensitive robotic systems for screening and diagnosis. These systems can improve education, social interactions, and resource access for autistic children despite challenges from limited understanding and inadequate specialized services. This study addresses these issues by designing emotionally expressive robots to support the developmental and educational needs of autistic children.

The second step focuses on creating more expressive gestures to enhance the gesture synthesis component. As the second step, this research centers on creating "Expressive Gestures" for the NAO robot, specifically designed to engage Sinhala-speaking children aged 2 to 6 years in Sri Lanka. These gestures aim to express various emotions while aligning with Sri Lankan cultural values, making them relatable and comprehensible for young kids. By considering cultural relevance and the developmental needs of this age group, the study ensures that the gestures are intuitive and effective at capturing children's attention. This strategy fosters trust and comfort during interactions and evaluates behavioral and emotional reactions.

Children with Autism Spectrum Disorder (ASD) often struggle with turn-taking, joint attention, and imitation, which impacts their social skills, communication, and learning abilities. Research in robotics has increasingly focused on incorporating social behaviors into robots intended for various settings, including homes, offices, and nursing facilities. Numerous studies have evaluated the role of socially assistive robots in diagnosing and treating children with ASD [2]. These studies indicate that children with ASD tend to engage more positively with robots than with human interactions. Additionally, research by Ricks and Colton found that children with ASD relate better to humanoid robots than to non-humanoid ones [3].

Fig. 1. Robot-child interaction environment.

Fig. 1 illustrates an environment designed for robot-child interactions using NAO robots for Autism Spectrum Disorder (ASD) screening. This study develops a multimodal interaction system for speech-based autism intervention among Sinhala-speaking children in Sri Lanka, integrating speech recognition, dialogue management, voice synthesis, and gesture synthesis to create an immersive therapeutic setting [4]. Building on prior work, which introduced a basic prototype with two emotional gestures, this research expands the gesture synthesis component by creating culturally adaptive and expressive gestures, such as happiness, sadness, fear, anger, dancing, sitting, and speaking, tailored for Sinhala-speaking children. The gestures' validity and effectiveness were evaluated through recorded videos reviewed by academics, students, and parents in Sri Lanka. This study's contribution lies in its focus on culturally relevant gestures, enhancing the system's applicability and relevance in therapeutic contexts.

This study addresses a critical gap in human-robot interaction for ASD therapy by designing culturally adaptive and emotionally expressive gestures for the NAO robot, tailored for Sinhala-speaking children. The research advances inclusive therapeutic tools by validating gestures through simulations, real-robot testing, and participant evaluations, demonstrating how culturally relevant gestures enhance child-robot interactions and emotional engagement.

II. Literature Review

Several key areas have been explored to advance socially intelligent humanoid robots. Emotional gestures are vital for enhancing communication and engagement in human-robot interaction. The need for culturally relevant design is highlighted to ensure gestures resonate with diverse users. Advances in gesture recognition and generation, incorporating machine learning for precise and adaptive motion, are discussed. Robots' ability to convey human-like emotions is also examined. Machine learning applications in motion generation and social learning enable robots to adapt to dynamic environments. Finally, integrating gestures with speech is crucial for natural and engaging interactions.

Study [5] focuses on using emotional expression in robotics, mainly on how the Nao robot can improve interaction between humans and robots. It urges people that robots need to express emotion to interact better and engage with humans. Finally, it details the selection and evaluation procedure of such human emotional postures and how they can be adapted to robotic forms, with external observers to assess the results. Researchers have successfully developed emotional postures for anger, sadness, and happiness, with happiness postures demonstrating the highest recognition rates.

Human-robot interaction (HRI) research underscores the importance of gestures in shaping human perceptions during joint tasks. Non-verbal communication, particularly gestures, is critical in HRI. The ACE framework studies multimodal interactions that enhance robot perception and evaluations. Robots using multimodal gestures consistently receive higher ratings than those relying solely on verbal communication, even when gestures and speech are incongruent, highlighting human perception's resilience. The study calls for flexible robot control architectures to optimize speech-gesture integration and advocates multimodal strategies to improve robotic systems. Future research should prioritize gesture-speech congruence to enhance HRI further [6].

Study [7] highlights the significant impact of robot gestures on human-robot interaction (HRI). It emphasizes that gestures, as a vital non-verbal communication tool, enhance interaction quality when combined with speech, fostering rapport and trust. Research indicates that multimodal communication leads to more positive evaluations of robots. Interestingly, even incongruent gestures—those that do not align with verbal messages—are still received favorably, suggesting that the presence of gestures can enhance perceptions of robots. The review also discusses the importance of flexible robot control architectures, which enable real-time adaptation of gestures and speech, further improving HRI. Overall, the findings underscore the critical role of gestures in shaping human perceptions and the need for continued exploration in this area to enhance interaction effectiveness.

Non-verbal communication is essential for enhancing human-robot interactions as robots become increasingly common. This review highlights behavioral overlays and proxemics as key strategies for improving robots' expressiveness and social awareness. The behavioral overlay model integrates non-verbal cues, enabling robots to adapt behaviors to interaction contexts and emotional states. The research underscores the value of individual responsiveness and interpersonal memory in tailoring robot responses and improving user experiences. The review emphasizes refining these models and addressing cultural differences to foster intuitive and socially aware robotic systems [8].

Study [9] explores how teleoperation has progressed in robotics by allowing remote control via human motion imitation, exemplified by the NAO robot, which utilizes the Kinect v2.0 sensor and a Hidden Markov Model (HMM) for motion recognition. Analytical geometry and vector algebra facilitate the calculation of joint angles for precise action replication. This focuses on a data augmentation method that enhances model training, establishing a framework for Kinect-based teleoperation of the NAO robot. Results indicate that this framework improves the robot's ability to mimic human motions and increases training efficiency and performance through data augmentation. The study underscores the potential of combining motion capture technologies with machine learning to advance robotic teleoperation systems.

Another study focuses on human-robot interaction using Indian Sign Language (ISL) to enhance communication with hearing-impaired individuals. It utilizes wavelet descriptors for feature extraction and possibility theory for gesture classification, achieving 92% accuracy for continuous ISL gestures, outperforming traditional models by 10%. The framework facilitates gesture recognition and text conversion for the NAO humanoid robot. However, challenges like high computational demands and sensitivity to lighting conditions remain, prompting further research. The study aims to improve ISL gesture recognition and address preprocessing and coarticulation detection issues, contributing to advancements in assistive technology with future efforts to increase accuracy and efficiency [10].

The emotional expressiveness of robots is vital for effective human-robot interaction, especially in social contexts. NAO, a humanoid robot, has been noted for its applications but lacks sufficient facial expressiveness. To address this, researchers developed pluggable eyebrows to enhance NAO's emotional conveyance, allowing for better expressions of anger and sadness. The mechanical design enables dynamic eyebrow movement, which was experimentally validated to improve emotional expression significantly. Findings indicate a linear correlation between eyebrow angle and expressed emotions. This advancement enhances NAO's interaction capabilities and underscores the importance of emotional expressiveness in human-robot relationships, paving the way for further research into socially interactive robots [11].

Study [12] emphasizes the importance of bodily expression of affect in human-robot interaction, noting that robots convey mood through bodily cues to enhance emotional and functional engagement. This integration allows robots to express emotions during tasks, improving user interactions. A parameterized behavior model was developed to represent mood states and their expressions, with studies showing that participants can recognize these mood expressions in humanoid robots without context. The findings suggest that practical mood expression enhances human-robot interactions, positioning robots as emotional companions or assistants, and highlight the need for further research into affective expression to create emotionally intelligent systems.

Advancements in humanoid robotics focusing on human-like movement generation. Early methods relied on motion capture and predefined libraries, but recent approaches utilize machine learning and probabilistic models. Movement primitives are flexible building blocks for complex behaviors, allowing robots to learn from human demonstrations. Bayesian networks enhance decision-making under uncertainty, improving real-time motion generation. Coupling neural networks enable precise arm movements and optimize motion-decision algorithms. Simulations, such as those with the NAO robot, validate these methods, indicating potential applications in assistive robotics, rehabilitation, and entertainment. Overall, integrating these strategies is expected to improve the realism and functionality of humanoid robots, fostering better human-robot interactions [13].

The study [14] goes through designing, implementing, and analyzing socially intelligent robots that can hold natural human-like conversations. It addresses the challenge of understanding human personality traits to advise robots to communicate better with the humans they interact with. Methodologies enabling speech translation, text adaptation, gesture generation, and SVM training to enhance interaction accuracy are covered. As discussed in the review, it is essential to model robot behavior on human personality traits and combine gestures with speech to be more engaging. Ultimately, it argues.

The integration of object recognition, speech recognition, and syntactic interpretation to enable robots to recognize and name objects via speech, leveraging advancements in RGB-D sensors for improved accuracy. Interactive communication with humans facilitates vocabulary acquisition, enhancing learning environments, while the convergence of these technologies underscores the importance of social learning in effective human-robot interaction (HRI). Future research should focus on refining these systems to expand their applicability, enabling more sophisticated and context-aware robot interactions [15].

III. Methodology

Fig. 2 shows a system overview that enables the NAO robot to perform synchronized gestures and speech using dialog management, ALMotion, ALLeds, and ALAudio modules. Multithreading ensures seamless interaction while sensor monitoring and posture control maintain safety. This integration creates natural, expressive behaviors for engaging and responsive robot-child interactions.

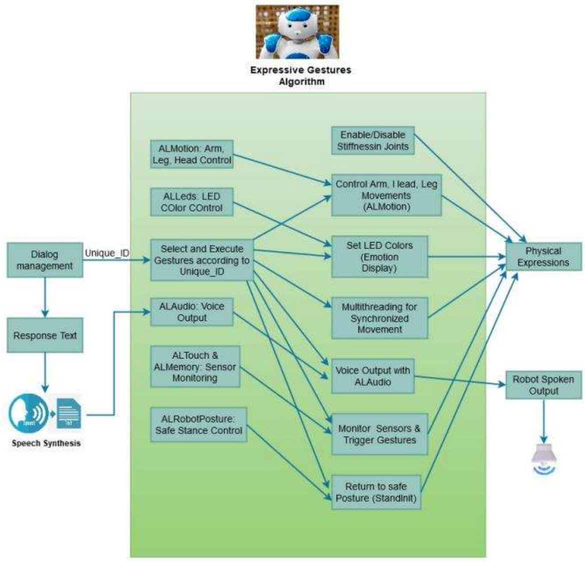

Fig. 2. System overview.

1) *Motion Generation for NAO Robot*

Without a pre-existing motion generation dataset for NAO, our approach involves creating and validating gestures using the Choregraphe simulator provided by SoftBank Europe. The Choregraphe software enables the design and testing of motion sequences for the NAO robot in a virtual environment. These motions were initially generated on the virtual robot within Choregraphe to minimize risks and errors before deployment. Once refined, the gestures were tested on a real NAO robot to ensure smooth execution.

To assess the expressiveness and appropriateness of these gestures, we validated them by demonstrating the robot's behavior to multiple observers. Feedback from individuals was collected to determine whether the gestures were engaging and suitable for conversations with children. This iterative process ensured the gestures' alignment with the intended emotional expressions.

2) *System Development Process*

The methodology consists of five key stages:

a) *Module Import and Proxy Establishment*

Python scripts using the NAOqi framework were developed to control the NAO robot. Key modules include ALProxy, for connecting to NAO-specific proxies; ALRobotPosture, for managing safe posture transitions; ALLeds, for controlling LED colors to express emotions visually; ALTextToSpeech, for generating speech to enhance gestures; ALMotion, for precise limb control using joint angles; and ALTouch/ALMemory, for monitoring touch interactions and enabling dynamic responses. These integrated proxies ensured seamless communication with NAO’s hardware, forming a robust foundation for motion generation and interaction design.

b) *Gesture Execution Based on Commands*

The system interprets high-level gesture commands representing emotions such as happiness, sadness, anger, fear, wave side dance, sitting rest, and free speech. These commands act as triggers for corresponding predefined gesture functions. Command Reception: the system receives an emotion command (e.g., “happiness”). Gesture Execution based on the command, the system calls the relevant gesture function (e.g., happiness_gesture, sadness_gesture), which combines movements, LED effects, and speech synthesis to express the desired emotion. Dynamic Adaptation for enhanced interaction, commands can also adjust specific gesture parameters like movement intensity or duration. This modular approach simplifies gesture management and facilitates future extensions to include new emotional expressions or behaviors.

c) Gesture Function Design

Each gesture function consists of multimodal components designed to work in harmony to express an emotion effectively. These components include. Arm Movements Arm positions and motions are achieved using the setAngles() method. For instance, raising arms or waving can indicate happiness or excitement. Head Movements Reflective movements, such as nodding or tilting the head, are implemented using the same method to add natural expressiveness. Leg Movements Adjustments to hip or knee angles contribute to body posture, such as crouching slightly for fear or bending for emphasis. LED Control LED colors and patterns are changed using fadeRGB() to signify emotions visually. For example, red LEDs can represent anger, while blue LEDs indicate sadness. Voice Output Speech is generated using tts_proxy.say() to complement gestures. For example, NAO might say, "I am so happy to see you!" during a happy gesture. By combining these components, the robot delivers synchronized and meaningful emotional expressions.

d) Multithreading for Gesture Synchronization

Multithreading was employed to achieve the natural and smooth execution of gestures. Multiple threads enable the simultaneous execution of Arm, leg, and head movements, LED color changes, and Speech output.

Each thread manages a specific aspect of the gesture, ensuring that movements, visual effects, and speech occur in parallel. For example, while the arms move to a happy position, the LEDs change to green, and speech output plays simultaneously. This level of synchronization is essential for producing cohesive and realistic gestures.

e) Integration and Development Tools

The developed system integrates with NAO using the NAOqi framework, enabling precise control and coordination of all robot functions. Key tools and technologies include NAO Robot V6, a Hardware platform for motion execution and testing. The NAOqi Framework is Middleware that provides APIs for controlling NAO's hardware components. The Choregraphe SDK Software suite is used for motion design, simulation, and deployment. Python (2.7, 3.3, and 3.12) is a Programming language for scripting gestures and managing interactions.

Integrating threading and subprocesses libraries to manage multithreading and inter-process communication ensures a robust, modular, and extensible system for motion generation and execution.

3) System Development Stages

Research focuses on progressively enhancing NAO's emotional interactions in conversations with children. The development process is structured into the following stages.

Typical Child – Limited Gestures Design and validate basic gestures for primary emotional states like happiness, sadness, and anger. Focus on simplicity to ensure reliability and appropriateness for initial testing.

Attention Management Develop gestures and behaviors that attract and maintain a child's attention, such as waving, head tilting, or dynamic LED patterns. Integrate speech output to complement attention-seeking gestures.

Face Tracking Implement face-tracking mechanisms to monitor the child's movements and adapt NAO's head and eye positions accordingly. This will enhance interaction by making NAO appear more responsive and attentive.

The current research stage focuses on developing and validating basic emotional gestures. Future stages will expand the robot's capabilities to include dynamic interactions, complex conversations, and adaptive responses based on real-time inputs. By following this structured methodology, we aim to create a robust system that enables NAO to deliver engaging and meaningful interactions with children.

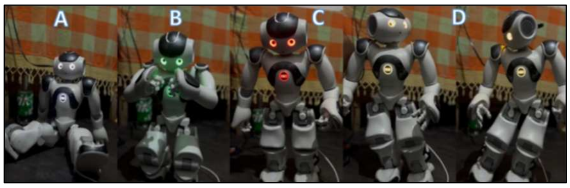

Fig. 3. Examples of gestures: sit down (A), fear (B), anger (C), and wave sides like dance (D).

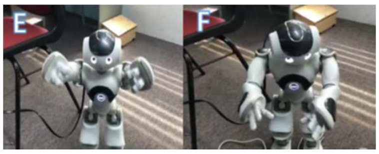

Fig. 4. Examples of gestures: happy (E), sad (F).

The robot's gesture postures showcased in Fig. 2. illustrate four postures: sitting down (A), displaying fear (B), expressing anger (C), and waving side to side like in a dance (D). Fig. 3. highlights two additional expressions: happiness (E) and sadness (F). These figures emphasize the robot's expressive capabilities.

The study utilizes a mixed-methods approach, merging qualitative and quantitative techniques to holistically assess the effectiveness of culturally adaptive gestures in robot-child interactions. By integrating qualitative insights from Screening and observations with quantitative data from surveys and performance metrics, the research provides a balanced evaluation of technical functionality and emotional resonance, aligning with established human-robot interaction frameworks to address culturally sensitive interventions effectively.

IV. Results and Evaluation

TABLE I provides a detailed breakdown of robotic gestures, including the specific movements of the head, hands, and legs and the timing required for each gesture. This information is crucial for programming the robot to execute these gestures accurately and seamlessly, enhancing its ability to interact with humans naturally and engagingly.

Using Google Forms, we evaluated the NAO robot's ability to recognize gestures and express emotions. These gestures included happiness, sadness, fear, anger, a side wave mimicking dance, a resting posture, and free speech. Short video clips of the NAO robot performing the respective action were embedded in each gesture. Respondents were asked to rate each gesture on a 5-point Likert scale based on clarity, emotional accuracy, and overall engagement.

TABLE I. GESTURES MOTOR VALUES AND TIMING

Gesture	Head (Pitch, Yaw)	Hands (Shoulder, Elbow, Wrist)	Legs (Hips, Knees)	LED Behavior	Timing (seconds)
Happiness Gesture	Pitch: [-0.2, 0.2]	Shoulder Pitch: [-1.0, 0.0]	Knee: [0.4, 0.0]	LEDs cycle through 5 colors	Head nod: 4, Hands: 2.5; Bounce: 0.3 down, 0.3 up per cycle. Total: ~5 seconds.
	Yaw: Static	Shoulder Roll: [0.2, -0.2]			
Sadness Gesture	Pitch: 0.5	Shoulder Pitch: [0.3]	-	Dim eye LEDs	Arm movement: 3 cycles of 1 second; Head shaking: 3 cycles of 0.5 seconds. Total: ~6 seconds.
	Yaw: [-0.5, 0.5]	Elbow Roll: [-2.0, 2.0]			
Anger Gesture	Pitch: [0.2]	Shoulder Pitch: [1.5, 1.5]	Knee: [0.3, 0.3]	Red eye and chest LEDs	Arms aggressively moved back: 2 seconds; Legs and body tilting: 1 second; LEDs: red during the gesture (~5 seconds).
	Yaw: Static	Elbow Roll: [-1.0, -1.0]	Hip: [0.2 tilt]		
Sit Down Posture	Pitch: 0	Shoulder Pitch: Static [-0.8]	Knee: [1.0]	-	Sit-down movement: 3 seconds. Static posture once seated.
	Yaw: Static	Elbow Roll: Static [0.5]			
Fear Gesture	Pitch: [0.5] (lowered)	Shoulder Pitch: [0.5, 0.5]	Static	Green eye and chest LEDs	Lower head: ~1 second; Arms raised to head: ~2 seconds; Crouch: ~3 seconds; LEDs: green. Total: ~6 seconds.
	Yaw: [0.0] (static)	Elbow Roll: [-2.0, 2.0]			
Wave Side	Pitch: Static, Yaw: [0.98, -0.98]	Shoulder Roll: [0.8, -0.8]	Hip Roll: [0.51, -0.51]	Yellow LEDs during wave	Wave to one side: 2.0 seconds; Reset: 0.5 seconds; Wave to the opposite side: 2.0 seconds; Total: ~6 seconds.
		Elbow Roll: [-0.59, 0.59]			
Speech Gesture	Pitch: -0.2, Yaw: [0.3, -0.3]	Shoulder Pitch: [0.5, 0.5], Elbow Roll: [0.3, -0.3]	Static	Blue LEDs for eyes during head movement	Initial position: ~0.5 seconds; Repeated hand movement (discussion): ~3 cycles (0.3 each); Total: ~4 seconds.

The form was distributed to diverse participants, including academics from universities across Sri Lanka and professionals from various industries. Our survey collected 122 responses, including 83 unmarried and 39 married respondents. The survey sample spanned multiple age groups and professional designations, such as assistant directors, engineers, researchers, and university students. Based on the collected data, it was possible to determine how effectively the robot conveyed each emotion and how expressive it was perceived overall. These findings were analyzed to refine the robot's gesture algorithms and improve its capability to interact with humans intuitively. With this participatory approach, every individual's perspective was included in the evaluation, allowing us to understand better how well the robot communicates emotionally with humans.

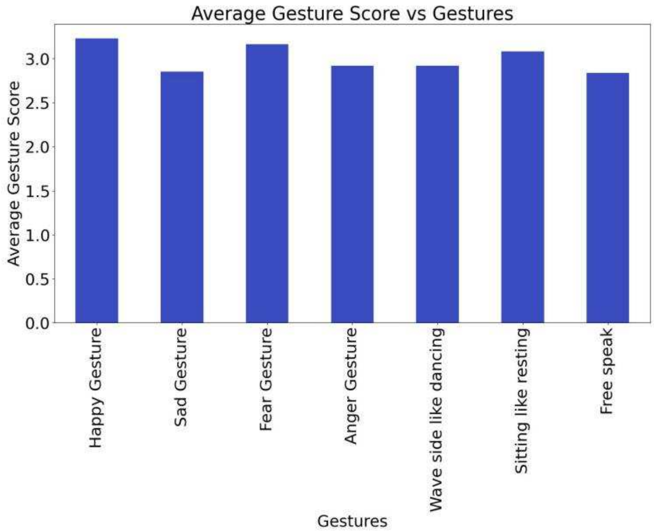

Fig. 5. Average gesture score vs gestures.

Fig. 5. illustrates the average scores for each gesture across the entire dataset. The bar chart compares how participants rated various gestures, such as Happy Gestures, Sad Gestures, Fear Gestures, and others, on a numerical scale. The results reveal which gestures were most rated as expressive or impactful, highlighting general trends in perceived emotional expressiveness. For instance, gestures like wave sides, such as dancing, fear gestures, and happy gestures, might show higher average scores, suggesting their effectiveness in engaging participants. In comparison, others, like sad Gestures and free-speech gestures, may exhibit comparatively lower scores, indicating areas for potential improvement or less impactful execution.

Fig. 6. explores the relationship between marital status and gesture perception by comparing the average gesture scores of married and unmarried participants. This grouped bar chart demonstrates how marital status influences participants' ratings of various gestures. Patterns may emerge, such as married individuals perceiving specific gestures (e.g., Sitting like resting or Anger Gesture) differently than their unmarried counterparts. The analysis helps understand demographic influences on emotional gesture interpretation and may guide adjustments for audience-specific robot interactions.

Fig. 7. focuses on the impact of age on gesture perception by categorizing participants into defined age ranges (e.g., 18-22, 22-26, ..., 54-58) and comparing their average gesture ratings. With color-coded gestures, this bar chart highlights variations in how different age groups perceive the robot's emotional expressions. For example, younger age groups might rate Happy Gesture and Wave sides like dancing higher, reflecting a preference for dynamic and expressive movements. In contrast, older age groups might favor calmer gestures like Sitting or resting. This age-based analysis provides insights into tailoring robot gestures to diverse audiences for enhanced interaction effectiveness.

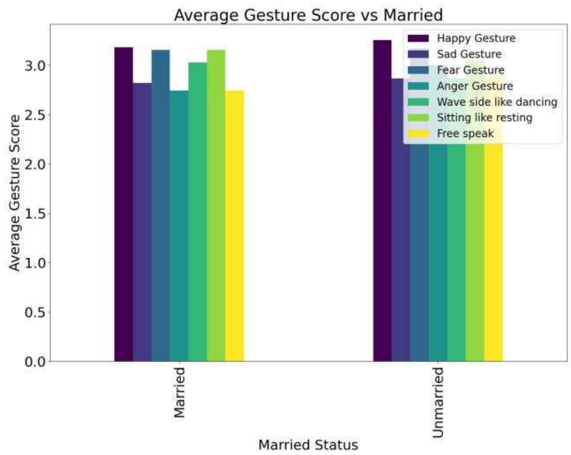

Fig. 6. Average gesture score vs married.

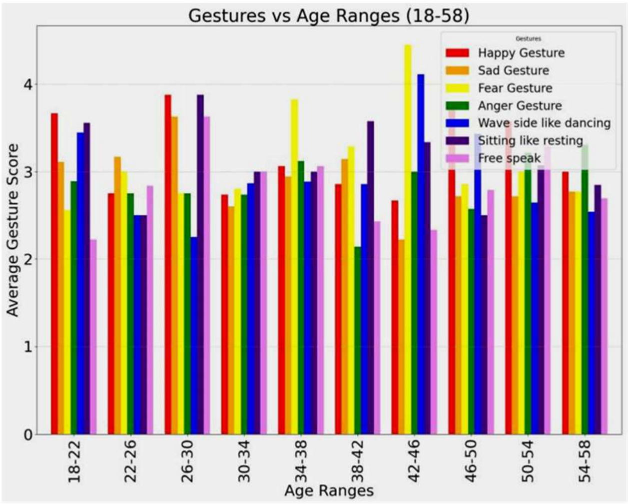

Fig. 7. Average gesture score vs age ranges (18-58).

V. CONCLUSION

This study designs and evaluates expressive gestures for the NAO6 robot to foster engaging interactions with children with ASD. Incorporating culturally relevant gestures enhances the robot's connection with its target audience while addressing developmental and social challenges. Simulations and real-world testing show improved engagement and emotional expression.

Results highlight the potential of socially assistive robots like NAO to enhance communication, interaction, and emotional recognition. Feedback from diverse participants, including university students and parents, confirms the robot's ability to convey emotions through synchronized movements and contextual responses. Gesture ratings offer insights into preferences influenced by demographic factors, enabling further personalization and adaptability.

Key contributions include validated motion sequences, guidelines for culturally appropriate interactions, and a framework for synchronizing gestures and speech using linguistic and contextual inputs. These findings have broad implications for healthcare, education, and therapy for children with ASD.

Future research should explore diverse datasets, cultural contexts, and long-term studies to assess interaction impacts. Integrating machine learning and real-time adaptability could further enhance the robot's responsiveness. This study emphasizes social robotics' critical role in addressing human needs and providing scalable solutions for fostering social and emotional development in children.

REFERENCES

[1] B. Craenen, A. Deshmukh, M.E. Foster, A. Vinciarelli, Shaping Gestures to Shape Personalities: The Relationship Between Gesture Parameters, Attributed Personality Traits and Godspeed Scores, in: 2018: pp. 699–704. https://doi.org/10.1109/ROMAN.2018.8525739.

[2] Kathleen Richardson, Mark Coeckelbergh, et al. "Robot Enhanced Therapy for Children with Autism (DREAM): A social model of Autism". In: IEEE Technology and Society Magazine (2018), pp. 30–39.

[3] D.J. Ricks and M.B. Colton. "Trends and considerations in robot-assisted autism therapy". In: 2010 IEEE International Conference on Robotics and Automation. IEEE. 2010, pp. 4354–4359.

[4] Jayawardena, C. S. Manukalpa, H. Bopage, P. Panduwawala, K. Pulasinghe and S. Rajapakshe, "A Multimodal Interaction System for Speech-Based Autism Intervention in Sinhala-Speaking Sri Lankan Children using the NAO Robot," 2024 6th International Conference on Advancements in Computing (ICAC), Colombo, Sri Lanka, 2024, pp. 205-210, doi: 10.1109/ICAC64487.2024.10851032.

[5] Erden, M.S. Emotional Postures for the Humanoid-Robot Nao. Int J of Soc Robotics 5, 441–456 (2013). https://doi.org/10.1007/s12369-013-0200-4

[6] Maha, Salem., Stefan, Kopp., Ipke, Wachsmuth., Katharina, J., Rohlfing., Frank, Joublin. "Generation and Evaluation of Communicative Robot Gesture." International Journal of Social Robotics, 4 (2012).:201-217. doi: 10.1007/S12369-011-0124-9

[7] Salem, Maha, et al. "Generation and evaluation of communicative robot gesture." International Journal of Social Robotics 4 (2012): 201-217.

[8] Brooks, Andrew G., and Ronald C. Arkin. "Behavioral overlays for non-verbal communication expression on a humanoid robot." Autonomous robots 22 (2007): 55-74.

[9] Balmik, Archana, Mrityunjay Jha, and Anup Nandy. "Nao robot teleoperation with human motion recognition." Arabian Journal for Science and Engineering 47.2 (2022): 1137-1146.

[10] Baranwal, N., Singh, A.K. & Nandi, G.C. Development of a Framework for Human–Robot interactions with Indian Sign Language Using Possibility Theory. Int J of Soc Robotics 9, 563–574 (2017). https://doi.org/10.1007/s12369-017-0412-0

[11] Albert, De, Beir., Hoang-Long, Cao., Pablo, Gómez, Esteban., Greet, Van, de, Perre., Dirk, Lefeber., Bram, Vanderborght. (2016). Enhancing Emotional Facial Expressiveness on NAO: A Case Study Using Pluggable Eyebrows. International Journal of Social Robotics, 8(4):513-521. doi: 10.1007/S12369-016-0363-X

[12] Junchao, Xu. "Body language of humanoid robots for mood expression." null (2014).:1711-1712. doi: 10.5555/2615731.2616139

[13] Yuan, Wei., Jing, Zhao. (2020). Designing Human-like Behaviors for Anthropomorphic Arm in Humanoid Robot NAO. Robotica, 38(7), 1205-1226. Available from: 10.1017/S026357471900136X

[14] Aly, A., Tapus, A. Towards an intelligent system for generating an adapted verbal and nonverbal combined behavior in human–robot interaction. Auton Robot 40, 193–209 (2016). https://doi.org/10.1007/s10514-015-9444-1

[15] Hinaut, Xavier, et al. "Humanoidly Speaking-How the Nao humanoid robot can learn the name of objects and interact with them through common speech." International Joint Conference on Artificial Neural Networks–IJCAI, Video Competition. 2015.