

Investigating the role of different social cues in the human perception of a social robotic arm.

Carlo La Viola^{1,*}, Laura Fiorini^{1,2}, Gianmaria Mancioffi¹ and Filippo Cavallo^{1,2}

¹*Department of Industrial Engineering (DIEF), University of Florence, Via Santa Marta 3, Firenze, Italy*

²*The BioRobotics Institute, Scuola Superiore Sant'Anna, Viale Rinaldo Piaggio 34, 56025 Pontedera (PI), Italy*

Abstract

The use of robotic arms in the rehabilitation context is increasing, leveraging the robotic capabilities related to object manipulation. Similarly, social robots are used in rehabilitation to provide features of social interaction with people. Even though these two robotic aspects are gaining visibility, the combination of both manipulation and social capabilities is often neglected. This work aims at filling this gap, introducing both manipulation and social capabilities in a scenario of human-robot interaction for rehabilitation. This work will define the best social configuration for a robotic arm, in terms of social cues. In particular, sound and expressive eyes cues will be linked to social movements designed using Laban Movement Analysis. Various combinations of movements and social cues will be combined and embedded in a robotic arm. Then 15 participants are recruited and asked to perform a human-robot interaction task (handover) with the robotic arm. A questionnaire is used to evaluate the user impressions in terms of Valence, Arousal, the Godspeed components of Animacy and Safety, and two custom questions related to movement fluidity and likeability. The results show that the best combination for this exercise is the one where both sound and eyes are present, even though the data show high scores of Arousal and Animacy for the configuration composed of only sound and movement.

Keywords

Social robots, social cues, Laban movement, HRI

1. Introduction

The robotic world is advancing towards technological improvements and the application of robots is spreading to various fields. In fact, it is possible to assist in the advancement of Social Robots [1] in the last years. Using social robots is relevant for rehabilitation purposes, as explained in [2] where the social robot is used to do rehabilitation sessions with children affected by autism, or in [3] where social robots are used to provide customized help to users suffering from cognitive changes related to aging and/or Alzheimer's disease. Many authors have investigated what takes for a robot to be social, and their works are summarized in the review by [4]. Among the various elements that can help in the perception of social robots, there is the similarity to human behavior, and the use of natural interfaces (i.e. speech) for

ICSR'22: International Conference on Social Robotics, ALTRUIST Workshop, Florence, IT

*Corresponding author.

✉ carlo.laviola@unifi.it (C. La Viola); laura.fiorini@unifi.it (L. Fiorini); gianmaria.mancioffi@unifi.it (G. Mancioffi); filippo.cavallo@unifi.it (F. Cavallo)

🆔 0000-0003-1745-0684 (C. La Viola); 0000-0001-5784-3752 (L. Fiorini); 0000-0001-8109-7956 (G. Mancioffi); 0000-0001-7432-5033 (F. Cavallo)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

communication, making the robot more human-like and socially accepted. Other social cues that can be abduced in making a robot social are the availability of a face and eyes, as explained in [5] where different face colors in a robot were linked to different emotions, or as in [6] where the use of animated eyes put on a not-humanoid robot helped the users in perceiving it as more alive and friendly. Sound also is an element that can make a robot more social. In the work of [7] the use of different sounds in relation to expressive gestures is proved to change the feeling of the users, and also in [8] various sounds applied to various robot movements were confirmed to improve the perception of valence, warmth, and competence of the robot itself. Finally, movement and body language are of paramount importance in humans' everyday interaction and play a key role in how people relate to each other. For this reason, movement is another social element of robots and some authors have explored how robotic movements can influence people's perception [9], while others explored robot gestures capable of conveying emotions [10]. Among the various approaches available in the literature to model robot movements, this work makes use of Laban movements, as introduced by other authors such as [11] and [12], [13]. These authors use Laban Movement Analysis, to model robotic movements and achieve emotional communication through movement. The authors that make use of Laban Movements for social robotic use humanoids (i.e. NAO, Softbank Robotics) [2] or puppet-like robots [14]. The limitation that emerged from the literature, is that such social capabilities are rarely combined with robotic manipulation, having normally only one of the two involved in the rehabilitation session and limiting the possibilities of the interaction (i.e. only social or only manipulation). Furthermore, there was no evidence in the literature of authors using a robotic arm, while in this work the robot used is indeed a robotic arm (Panda, Franka Emika, Germany). The authors of this study have already explored Laban movements on a robotic arm [15], suggesting that it can trigger an emotional feeling in the users; the use of a robotic arm is also effective in performing exercises for people who have to rehabilitate upper limbs [16]. The novelty of this work is to present a new concept of rehabilitation robotics where physical and cognitive rehabilitation is achieved using a social robotic arm capable of manipulation and social behavior. The implementation of various social cues (i.e. face, sound, movement) on a robotic arm will be presented and they will be used in an HRI task. The analyses will converge on what is the impact of social cues on the human perception of the robot, and if some social cues are dominant over others. Moreover, the purpose of this work is to define what kind of Laban movements are the best in a scenario of robotic rehabilitation.

2. Materials and Methods

This study is based on the definition of an HRI task for the evaluation of the proposed novelty. The interaction task design is simple, yet relevant in the realm of HRI and is a starting point for future implementation of this system. A handover task will be presented, which is composed of multiple steps that can be used singularly or grouped in the execution of a rehabilitation exercise (e.g. tower of London [17]). In this section, the procedure to put the experiment in place is described and further details are provided regarding the setup.

Table 1

Summary of Laban parameters possible values

Laban component	Parameters name	Possible Values
Laban Effort Graph	Space	Direct, Indirect
	Time	Sudden, Sustained
	Flow	Bound, Free
	Weight	Strong, Light
Laban Space	Horizontal	Advancing, Retreating
	Vertical	Ascending, Descending
	Wheel	Spreading, Enclosing

Table 2

Values for the Direct and Indirect Movement

Movement ID	Laban Effort Graph (LEG)				Laban Space (LS)		
	Space	Time	Weight	Flow	Horizontal	Vertical	Wheel
Indirect (I)	Indirect	Sudden	Light	Free	Spreading	Ascending	Advancing
Direct (D)	Direct	Sudden	Strong	Bound	Spreading	Ascending	Advancing

2.1. Design of social cues

In the first place, this work focused on how to design the social cues that should be used during HRI, to achieve a particular task objective. As described in Section 1 this work uses 3 social cues. *I) Robotic Movement*: stemming from a previous work of the same authors [15], the movement design is based on Laban Movement Analysis (LMA). LMA was defined by Rudolph Laban in 1948 [18] as a system for describing, discussing, and documenting human body movements according to different components, namely Laban Effort Graph (LEG) referring to the body's inner use of energy, and Laban Space (LS), referring on how the body uses the surrounding space. The LEG component is composed of four parameters that can have one between 2 values and the LS component can vary according to the 3 parameters (corresponding to 3 planes of movement) that change between two opposite values. A schema of such a configuration and all the possible values are summarized in Table 1.

The different combinations of LEGs parameters and the LS parameters originate different movements, that are designed to convey emotional information. The set of movements to be used in this work comes from the previous work [15] where various Laban Movements were shown to a pool of 120 participants and their preferences were collected. The most preferred movements were the one modeled after the happiness emotion and the one modeled after the anger emotion; in this work, the happy movement will be called Indirect and the angry one will be called Direct. Their Laban configuration is presented in Table 2.

I) Eyes: the face was selected from another previous work [6]. This same face was used on a wheeled social robot deployed in the context of social interaction with older adults and disabled people. The expressivity of the eyes (which is composed of blinking and moving eyes) was liked by all the participants in the previous work and for this reason, was included as is in the study. *I) Nonverbal sound*: the inspiration for the sound generation was taken from [8], where the

authors generate different sounds for different robots. Specifically, for this work, the sound that most of the others guided the design was the one taken from their sonification of UR5 movements. The generation is done programmatically using the library `pygame` (<https://www.pygame.org/docs/>), and the sound is composed with an initial descending C Major arpeggio, followed by a G Maj chord (5th), then an F Maj chord (4th), to end with an ascending C Major arpeggio. Change of chords was triggered by events in the robot movement, namely ball picking, handover, start, and end of the interaction session.

2.2. Software system to support the experiment

The overall system is run on a Ubuntu OS and takes advantage of the backbone infrastructure provided by ROS [19]. Each of the social cues is written in a node and registered to a common topic, which is published by a */controller* that takes charge of the synchronization of the system. It is worth highlighting that even though the overall system is ROS based, the robotic arm movements were achieved thanks to `Frankx` [20] which is a movement library developed for the Franka Emika Panda robot and that allows control over velocity, acceleration, and jerk. The final part of the ROS system is a node for saving data after the interactions, which connects to the disk and saves all the interaction session data in a CSV format, for later analysis. The data saved are user feedback regarding the interaction session and general information on the robot, such as time duration of the movement, and information on which action it is performing. At last, a node that records bag data from a RealSense camera (Intel, USA) is synchronized with the system for autonomous recording and saving of the different sessions. The overall schema for all the ROS nodes is depicted in Fig. 1. Another element of the system, unrelated to ROS, is a Docker container that hosts a Flask app; such an app allows connection from external elements to the app that is used to administer the questionnaires to the participants in the experiment and allows to save the data on the disk passing by a ROS node, contacted using `RosBridge` [21]. All system is triggered by a ROS service which is manually called by the experimenter. This choice was done to allow full control over the session duration and manage any unexpected event that could have happened during the experimentation.

2.3. Participants

A total of 15 participants were enrolled for this test. The people enrolled were researchers, medical doctors, psychologists, and nurses that volunteered in the experiment. No eligibility criteria were selected and all volunteers were accepted into the study. The participants are composed of 60% male and 40% female, and their education level is composed of 3 groups (33.3% Bachelor's Degree, 53.3% Master's Degree, and 13.3% Ph.D.). The mean age of participants is 30.47 years, with a standard deviation of 4.64. Before starting, all the participants were informed of the test and it was given to them a formal description of the experiment on paper; all of them were asked to sign informed consent.

2.4. Experimental procedure

The task to be performed during this experiment is the handover of a colored ball between the robot and the participant, with the participant that places the ball in a dedicated box to end

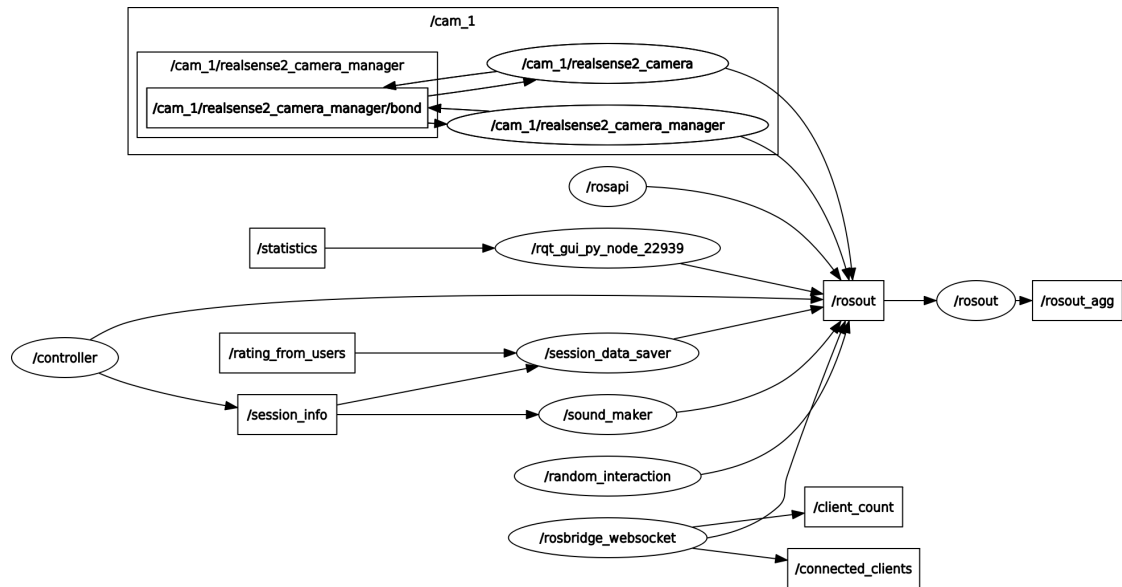


Figure 1: Custom ROS nodes schema.

the interaction. The social cues will be proposed in 4 possible combinations (only movement, movement and sound, movement and face, movement, face, and sound), coupled with the 2 movements (Direct or Indirect). Hence, there will be a total of 8 trials for each participant (4 social cues combinations by 2 movements). The experiment can be divided into 3 main phases, namely the robotic arm movement, the human-robot handover, and the rating of the interaction by the user. The robotic arm movement is in charge of picking up the colored ball and passing it to the user. The main design idea of the movement is introduced in the previous paragraph, and can be further described as a combination of 5 main blocks: i) The robotic arm turns toward the colored balls, ii) The robotic arm performs a movement to get closer to the balls and picks one of them, iii) The robotic arm turns toward the user, iv) The robotic arm performs a movement and gets closer to the user to handover the ball, v) The robot goes back to the idle position. The choice of which ball to pick, the order of movements, and the social cues combination were random and automatically decided by the system. After the robotic arm movement, the human-robot handover phase started, with the participant that grasps the ball offered by the robot, and puts it in a dedicated box with the same color as the ball displayed on top. Eventually, the participant is asked to complete a questionnaire to rate the experience of that specific session. The questionnaire used in this work was structured with a first part where it asked the user to rate the interaction according to the Self Assessment Manikin (SAM) [22], in terms of Valence, Arousal and Dominance. Then the user is asked to complete two elements of the Godspeed questionnaire [23], specifically the Animacy and the Safety component. At last two custom questions were asked, regarding the fluidity and the likeability of the movement. At the end of the 8 interactions, the experimenter asked 5 open questions to the participants, to acquire their qualitative feedback. Such questions were: 1) What do you think about the movements? 2) What do you think about the sound? 3) What do you think about the face? 4)

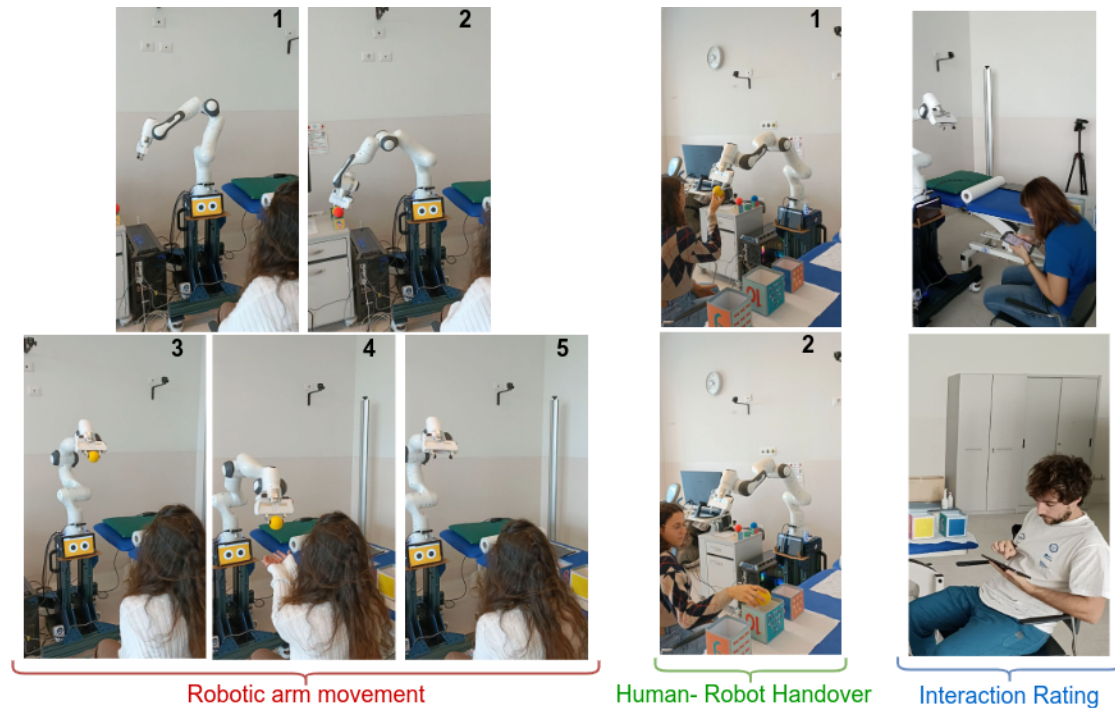


Figure 2: The 3 different phases of the experiment.

Did the movements convey to you any emotion or sensation? 5) Imagine bringing the robot home to use for your own purpose, which configuration (sound, face, movement) would you like to have? In this work only answers to question 5 will be analyzed. A schema of the phases for this experiment is available in Fig. 2.

2.5. Data Anlysis

The analysis focused on the data collected from the questionnaires. For the SAM scores, the values were extracted on a scale of 1 to 9; the Godspeed questions were summed up together to obtain an aggregated index with all the scores for the Godspeed elements where the maximum value is 25 for animacy and 15 for safety and the minimum values are respectively 5 and 3. Finally, for the likeability and fluidity scores, the value is taken and evaluated on a scale from 1 to 5. The means for each answer's group were evaluated for each interaction. The sessions reported in Table 3 are divided into columns by social cues (no social cues, eyes, sound, both sound and eyes) and by Direct (D) or Indirect (I) movement. The means extracted from the questionnaire were then compared across all the different iterations and the comparison between the different movements and social cues combination are analyzed and discussed. Due to the low number of subjects (15), only qualitative analyses were conducted and no statistical significance was looked for. Finally, the open answers to question 5 were analyzed using content analysis [24] and 2 main components were extracted, namely which configuration among the 4 presented was the one preferred and which movement was the most liked by participants.

Table 3

SAM, Godspeed and Custom aggregated values for each session

		NO SOCIAL CUES		EYES		SOUND		SOUND AND EYES	
		D	I	D	I	D	I	D	I
SAM	Valence Arousal	6.56	6.31	6.74	6.34	6.58	6.41	7.01	7.19
		5.14	5.38	5.09	4.88	4.86	5.43	5.11	5.07
Godspeed	Animacy Safety	15.67	16.60	16.33	16.73	16.53	17.67	17.67	17.27
		11.80	12.13	12.00	12.00	11.53	12.33	12.53	12.33
Custom	Likeability Fluidity	3.87	4.04	3.74	3.74	3.91	4.06	4.07	4.13
		3.62	3.92	3.22	3.56	3.60	3.77	3.79	4.05

3. Results

In the first place, the mean values for the SAM, Godspeed, and Custom values were computed. The complete results are depicted in Table 3. The Valence has the highest values for the session with all the social cues and the Indirect movement, asserting that this interaction is the most positive. On the contrary, Arousal has the maximum value in correspondence to the session with only sound and Indirect movement. So the movement that maximizes the Arousal value is the Indirect one; indeed the social cues play a role in the perception, and the session which has the highest Arousal score is the one with sound only.

As for the Godspeed, it is possible to identify the highest Animacy scoring in the sessions with only sound and with sound and eyes, for Indirect and Direct movement respectively. The sound seems an element that for the Indirect movement is capable of increasing the perception of Animacy (17.67 against 16.53 for Direct movement). Regarding Safety, it is possible to notice how the highest value is in the last session for the Direct movement. At a first glance, it is possible to say that the Godspeed parameters used in this study are maximized for Direct movement. Finally, the highest values for both likeability and fluidity are in the session with all social cues and Indirect movement. When looking at the open answers the results partially reflect the values highlighted in the tables. The preferred movement (Figure 3a) is the Indirect one (53.3%), but the 33.3% of people did not express a preference; only 13.3% rated the Direct as the preferred movement. For the configuration (Figure 3b), the preferred configuration of social cues is the one where both sound and eyes are present (60%), while 26.7% of participants liked only eyes and the 13.3% liked only the sound.

4. Discussion

The results reported in Table 3 show that the configuration of the social cues with the highest scores is the one where eyes, sound, and movement are used at the same time. Including only the eyes or the sound in the session does not seem to affect significantly the perception, even though it is worth mentioning that the use of only sound has been rated at the highest for the

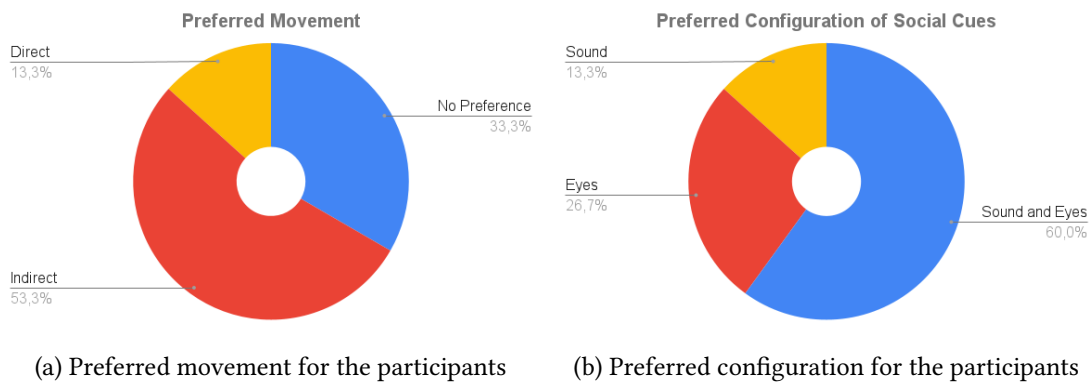


Figure 3: Pie charts for the preferences related to movement and configuration

SAM component of Arousal and for the Godspeed part of Animacy. This evidence suggests that sound was more appreciated than the eyes in communicating social information. The reasons for this are mainly 2: the focus of the user was on the gripper part of the robotic arm, so the eyes, placed in the bottom part of the robot, had a less relevant impact on the perception. Then, being the gaze of the user focused on the robotic gripper, it was easier to feel a social cue that is perceived through another one of the 5 human senses, the hearing. When evaluating the open answer left by the participants, the results are not fully aligned with the quantitative data. In fact, even though the quantitative data for the Godspeed analysis showed an overall preference in terms of Safety and Animacy of Direct movement, the open answers show a preference for Indirect movement. Similarly, when asking about the preferred configuration, even though the quantitative data for the session involving only sound is higher, the preferences are lower for only sound than for only eyes. This point suggests that there is not a perfect alignment between the quantitative and qualitative data and this information should be furtherly investigated in future studies. The answers regarding the most preferred session are similar, being the session of both eyes and sound the best (60%) for this sample of participants. The results collected confirm the work from [8] and [6], attesting that sound and face put on a robot increase its social perception from humans. Moreover, this work goes beyond such results introducing social movements, and indicating that robotic arms can be perceived as social and socially active if they respect some standards of movement and are equipped with both eyes and sound social features. The results suggest also an advancement with respect to the previous work [15], establishing what movement is the most appreciated. Looking at Table 3 it can be evinced how the highest values are (all but Safety) in the column related to the Indirect movement. This will guide further implementation of the system, including movements that are more indirect for implementation in real rehabilitation applications.

5. Conclusion

The aim of this work was to evaluate the most preferred combination of robotic arm movement and social cues. Secondly, there was the purpose to understand what kind of movement is

preferred by people. The preferred configuration has been verified to be the one with sound and eyes, and the preferred movement, regardless of social cues, is the Indirect one. Further evolution of this study will aim at enlarging the sample size to perform also statistical analysis of the results. Moreover, it will be possible to verify the correlation between perception and users' personalities. Nonetheless, further experiments should validate the findings of this work and such findings will be used to perform a real rehabilitation exercise with healthy people, to evaluate their perception of the robot and the validity of the social interaction.

References

- [1] T. B. Sheridan, A review of recent research in social robotics, *Current opinion in psychology* 36 (2020) 7–12.
- [2] T. Zhu, Z. Xia, J. Dong, Q. Zhao, A sociable human-robot interaction scheme based on body emotion analysis, *International Journal of Control, Automation and Systems* 17 (2019) 474–485.
- [3] A. Tapus, C. Țăpuș, M. J. Matarić, The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia, 2009 IEEE International Conference on Rehabilitation Robotics, ICORR 2009 (2009) 924–929. doi:10.1109/ICORR.2009.5209501.
- [4] A. Henschel, G. Laban, E. S. Cross, What makes a robot social? a review of social robots from science fiction to a home or hospital near you, *Current Robotics Reports* 2 (2021) 9–19.
- [5] T. Ariyoshi, K. Nakadai, H. Tsujino, Effect of Facial Colors on Humanoids in Emotion Recognition Using Speech, Technical Report, 2004. doi:10.1109/ROMAN.2004.1374730.
- [6] A. Sorrentino, L. Fiorini, C. La Viola, F. Cavallo, Design and development of a social assistive robot for music and game activities : a case study in a residential facility for disabled people *, 2022, pp. 2905–2908.
- [7] E. Frid, R. Bresin, Perceptual Evaluation of Blended Sonification of Mechanical Robot Sounds Produced by Emotionally Expressive Gestures: Augmenting Consequential Sounds to Improve Non-verbal Robot Communication, *International Journal of Social Robotics* 14 (2022) 357–372. URL: <https://doi.org/10.1007/s12369-021-00788-4>. doi:10.1007/s12369-021-00788-4.
- [8] B. J. Zhang, N. Stargu, S. Brimhall, L. Chan, J. Fick, N. T. Fitter, Bringing WALL-E out of the Silver Screen: Understanding How Transformative Robot Sound Affects Human Perception, 2021 IEEE International Conference on Robotics and Automation (ICRA) (2021). doi:10.1109/ICRA48506.2021.9562082.
- [9] J. Xu, J. Broekens, K. Hindriks, M. A. Neerincx, Mood contagion of robot body language in human robot interaction, *Autonomous Agents and Multi-Agent Systems* 29 (2015) 1216–1248. doi:10.1007/s10458-015-9307-3.
- [10] J. Li, M. Chignell, S. Mizobuchi, M. Yasumura, Emotions and messages in simple robot gestures, volume 5611 LNCS, 2009. doi:10.1007/978-3-642-02577-8_36.
- [11] S. J. Burton, A. A. Samadani, R. Gorbet, D. Kulić, Laban movement analysis and affective

- movement generation for robots and other near-living creatures, in: Springer Tracts in Advanced Robotics, volume 111, 2016, pp. 25–48. doi:10.1007/978-3-319-25739-6_2.
- [12] H. Knight, R. Simmons, Expressive motion with x, y and theta: Laban Effort Features for mobile robots, in: IEEE RO-MAN 2014 - 23rd IEEE International Symposium on Robot and Human Interactive Communication: Human-Robot Co-Existence: Adaptive Interfaces and Systems for Daily Life, Therapy, Assistance and Socially Engaging Interactions, Institute of Electrical and Electronics Engineers Inc., 2014, pp. 267–273. doi:10.1109/ROMAN.2014.6926264.
- [13] H. Knight, R. Simmons, Laban head-motions convey robot state: A call for robot body language, in: Proceedings - IEEE International Conference on Robotics and Automation, volume 2016-June, Institute of Electrical and Electronics Engineers Inc., 2016, pp. 2881–2888. doi:10.1109/ICRA.2016.7487451.
- [14] C. Clark, L. Sliker, J. Sandstrum, B. Burne, V. Hagggett, C. Bodine, Development and preliminary investigation of a semiautonomous socially assistive robot (sar) designed to elicit communication, motor skills, emotion, and visual regard (engagement) from young children with complex cerebral palsy: A pilot comparative trial, *Advances in Human-Computer Interaction 2019* (2019).
- [15] C. La Viola, L. Fiorini, G. Mancioffi, J. Kim, F. Cavallo, Humans and Robotic Arm : Laban Movement Theory to create Emotional Connection *, in: *Ro-MAN*, 2022.
- [16] D. Eizicovits, Y. Edan, I. Tabak, S. Levy-Tzedek, Robotic gaming prototype for upper limb exercise: Effects of age and embodiment on user preferences and movement, *Restorative Neurology and Neuroscience* 36 (2018) 261–274. doi:10.3233/RNN-170802.
- [17] P. Anderson, V. Anderson, G. Lajoie, The tower of london test: Validation and standardization for pediatric populations, *The Clinical Neuropsychologist* 10 (1996) 54–65.
- [18] R. von Laban, *Modern educational dance.*, 1948. URL: <http://books.google.com/books?id=nsAKAQAIAAJ>.
- [19] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, A. Y. Ng, et al., Ros: an open-source robot operating system, in: *ICRA workshop on open source software*, volume 3, Kobe, Japan, 2009, p. 5.
- [20] L. Berscheid, T. Kröger, Jerk-limited Real-time Trajectory Generation with Arbitrary Target States (2021). URL: <https://github.com/pantor/ruckig>. arXiv:2105.04830v2.
- [21] R. Toris, J. Kammerl, D. V. Lu, J. Lee, O. C. Jenkins, S. Osentoski, M. Wills, S. Chernova, Robot web tools: Efficient messaging for cloud robotics, in: *2015 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, IEEE, 2015, pp. 4530–4537.
- [22] M. M. Bradley, P. J. Lang, Measuring emotion: The self-assessment manikin and the semantic differential, *Journal of Behavior Therapy and Experimental Psychiatry* 25 (1994) 49–59. URL: <https://www.sciencedirect.com/science/article/pii/0005791694900639>. doi:[https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9).
- [23] C. Bartneck, D. Kulić, E. Croft, S. Zoghbi, Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots, *International journal of social robotics* 1 (2009) 71–81.
- [24] K. Krippendorff, *Content analysis: An introduction to its methodology*, Sage publications, 2018.