Personalizing User Engagement Dynamics in a Non-Verbal Communication Game for Cerebral Palsy

Nathaniel Dennler¹, Catherine Yunis², Jonathan Realmuto³, Terence Sanger³, Stefanos Nikolaidis¹, and Maja Matarić¹

Abstract— Children and adults with cerebral palsy (CP) can have involuntary upper limb movements as a consequence of the symptoms that characterize their motor disability, leading to difficulties in communicating with caretakers and peers. We describe how a socially assistive robot may help individuals with CP to practice non-verbal communicative gestures using an active orthosis in a one-on-one number-guessing game. We performed a user study and data collection with participants with CP; we found that participants preferred an embodied robot over a screen-based agent, and we used the participant data to train personalized models of participant engagement dynamics that can be used to select personalized robot actions. Our work highlights the benefit of personalized models in the engagement of users with CP with a socially assistive robot and offers design insights for future work in this area.

I. INTRODUCTION

Cerebral palsy (CP) is one of the most prevalent motor disorders in children [1], affecting around 0.2%-0.3% of all live births in the United States [2]. The main symptom of CP is involuntary muscle contractions that lead to repetitive movements [3] which can greatly affect a child's ability to communicate with caregivers and peers [4]. This symptom necessitates the use of active orthoses to facilitate proactive communication and to aid in motor rehabilitation [5].

Retraining motor skills, however, requires repetitive and lengthy sessions to be effective [6]. In children especially, this can lead to disengagement with the therapeutic activity, negatively affecting functional outcomes. Thus, we aim to facilitate engaging therapeutic activities through the development of an engaging game between a participant and a socially assistive robot (SAR) [7] [8]. The robot encourages the participant to perform (and therefore practice) non-verbal communicative gestures while providing social reinforcement as the participant makes progress in the game.

This work explores the effect of both physical and social factors of the interaction design on the ability to effectively engage participants. Physically, we investigate the embodiment of the agent that delivers the game. Several recent works have described the positive effect of strongly embodied agents, such as robots, on the engagement of participants over weakly embodied agents, such as computers (see review by Deng et al. [9]). Socially, we investigate how the feedback provided by the agent throughout the game affects participant engagement. Understanding how robots can effect engagement dynamics is an under-explored area of human-robot interaction (HRI) [10].

Both physical and social factors are investigated through a user study of participants with CP. We found that participants preferred interacting with the SAR compared to a screenbased agent but did not observe any significant differences in engagement levels between the two conditions, which we attribute to individual differences in how participants responded to the robot's actions, as detailed in Section V-B. To explore user engagement further, we then developed a probabilistic model for personalizing the robot's actions based on an individual participant's responses to the robot, and show in simulation that this improves the users' engagement levels compared to models that are not personalized. Together, the results of this work indicate the promise of personalized SAR for helping individuals with cerebral palsy to practice non-verbal communication movements.

II. RELATED WORK

Engagement is a key factor in measuring the quality of HRI scenarios [10], and in particular has been studied extensively as means of maintaining user interest [11]. Userspecific perceptual models to identify user engagement have been explored in the context of the autism spectrum disorder [12]–[14], where user behavior varies significantly due to personal differences. These personal differences are also present in CP populations, where there is a great variance between individuals in how motor function is impacted.

Consequently, there has been an emphasis on developing personalized user models to facilitate interactions (see reviews by Clabaugh et al. [15] and Rossi et al. [16]). Personalized interactions have shown promising results in various domains, ranging from rehabilitation [17] to robot tutoring systems [18], by implementing robot action selection based on personalized user models; however, few have studied engagement dynamics.

A review of several studies [19] concludes that SARs have been effective for clinical populations diagnosed with CP, demonstrating that physical robots can elicit positive responses from users with CP who are performing repetitive physical exercise tasks. Robots as partners in game-like therapeutic physical activities have been shown to create engaging experiences for users and lead to increased motivation

¹Nathaniel Dennler, Stefanos Nikolaidis, and Maja Matarić are from the Computer Science Department at the University of Southern California in Los Angeles, CA {dennler, nikolaid, mataric} @usc.edu

²Catherine Yunis is from the Biomedical Engineering Department at the University of Southern California in Los Angeles, CA cyunis@usc.edu

³Jonathan Realmuto and Terence Sanger and are from the Electrical Engineering and Computer Science Department at University of California, Irvine in Irvine, CA {jrealmut, tsanger}@uci.edu





(b) Screen Condition

(c) Animated Face Used by Robot and Screen

(d) Orthosis

Fig. 1: Study setup. The participants used non-verbal gestures to communicate with the robot or computer screen, depending on the experiment condition. The external 3D camera collected real-time participant hand movement data used by the robot.



Fig. 2: Stages of the within-subject experiment design.

[20]. The importance of engagement is emphasized in studies involving movement exercises for CP, and quantitative measures of engagement are well-established for this context [21]. Given the success of using SARs with this population, we aim to understand how SARs can shape user engagement in practicing repetitive exercises.

III. METHODS

A. Study Setup

The study setup consisted of the participant sitting at a table and facing a computer screen or a robot, both at eyelevel, as shown in Figure 1. The experimenter was present in the room for safety and to collect verbal questionnaire data. Finally, the participant's parent was located in the hallway outside, seated separately and not interacting with the study.

We used the tabletop LuxAI QT robot [22], shown in Figure 1-a; the robot is 25 inches tall, has arms with three DOF arms and a head with two DOF with a screen face. The robot was modified to work with the CoRDial dialogue manager [23] that synchronizes facial expressions with text-to-speech.

The screen condition used a standard 19" monitor that displayed the same simple animated face as the robot's at the same size, as shown in Figure 1-c, and used the same dialogue manager. The robot and screen agent used the same voice, the same facial features, and the same facial expressions, as well as the same machine vision algorithm. The strongly-embodied robot moved and gestured in the shared desk space with the participant, while the weaklyembodied screen was stationary, as shown in Figure 1-b.

The participants wore an orthosis shown in Figure 1-d that used fabric-based helical actuators to support the participant's thumbs-up and thumbs-down gestures. The orthosis

was controlled by a Beaglebone microprocessor [24], actuated with a compressed air tank, and was attached to the participant with Velcro strips for facile donning and doffing [5]. The orthosis was worn throughout the session and was not a manipulated variable in the experiment.

The participant's thumb angle was measured by an RGBD camera and transmitted through a ROS network [25]. A webcam placed on the table in front of the participant captured and recorded the participant's facial expressions. An emergency stop button was provided to the participant for terminating the interaction at any point.

B. Interaction Design

At the start of the session, each participant demonstrated a thumbs-up and thumbs-down gesture to generate a baseline for their individual range of motion. Next, the robot explained the number-guessing game, telling the participant to think of a number between 1 and 50, and to communicate the number secretly to the experimenter, by whispering or typing the number on an iPad. At each turn of the game, the robot guessed the number and asked the participant if the guess was correct. The participant answered yes or no by making a thumbs-up or thumbs-down gesture, respectively, using the arm with the orthosis. If the robot guessed incorrectly, it asked if the number was higher than the guess. The participant then answered thumbs-up if the number was higher, and a thumbs-down if the number was lower. The robot ensured that the number of thumbs-up and thumbsdown gestures were approximately equal by tracking the number of each and guessing higher or lower than the target number to keep the counts balanced. The robot continued to guess numbers randomly from a range of decreasing size as it narrowed in on the correct answer. Once the robot correctly guessed the number and the participant signalled with a thumbs-up, the robot asked to play the game again. The participant responded with another thumbs-up or thumbsdown gesture.

Every time the participant used a thumbs-up or thumbsdown gesture to respond to the robot, the robot responded with feedback that combined verbal, physical, and facial action, based on the quality of the gesture and the history of the participant's gestures. Specifically, the feedback was a *clarifying* utterance, an *encouraging* utterance, or a *reward-ing* utterance, accompanied with a corresponding physical gesture and facial expression. *Clarifying* actions were given when the participant's response was not legible. *Encouraging* actions were given when the angle the participant's thumb made was near their personal baseline value. *Rewarding* actions were given when the participant's thumb angle exceeded their personal baseline value. All verbal, physical, and facial components of these feedback actions were selected randomly from a set of appropriate components for each action, to avoid repetition.

C. Study Design

The study used a within-subjects design shown in the block diagram in Figure 2; the participants interacted with the robot in a single session that lasted approximately one hour from the participants entering the room to their departure. The session was divided into four blocks, with periods of rest in between. The first three blocks lasted up to 10 minutes each and had the participant play as many games with the robot/screen as desired, while the final block was open-ended, with no fixed duration. Between blocks, the participant rested for at least one minute or until they were ready to move to the next block to mitigate effects of muscle fatigue. The first block served as a practice block to familiarize the participant with the interaction. In that block, the participant interacted with the robot while the orthosis was not powered and thus not assisting their movement. In the second block, the participant interacted with the robot with the orthosis powered on. After the second block, the experimenter verbally administered a questionnaire about the participant's experience with the robot. In the third block, the participant interacted with a computer screen with the orthosis powered on. After the third block, the experimenter verbally administered a questionnaire on the participant's experience with the screen-based agent. The fourth block was optional, and the participant was given a choice of playing with the robot, playing with the screen, or ending the session.

D. Hypotheses

Since strongly-embodied physical agents have been shown to increase engagement and positive outcomes in therapeutic tasks [9], [19], the following hypotheses were tested:

H1: Users with CP will **prefer** the robot over the screen. H2: Users with CP will be **more engaged** when interacting with the robot than the screen.

E. Study Population

We recruited 10 participants (3 female, 7 male) diagnosed with CP and having symptoms of dystonia in at least one upper limb. The age range of the participants was 9-22 years, with a median age of 15 years. The gender imbalance is representative of the higher prevalence of males in CP populations [26], and the large age range reflects the challenges of recruitment of this population. Half of the participants wore the orthosis on their left hand, and the other half wore the orthosis on their right hand. All participants

TABLE I: Survey questions and associated factors of Companionship (C), Perceived Enjoyment (PE), and Perceived Ease of Use (PEU).

Survey Question	Factor
How much do you like playing with the {robot, screen}?	C [27]
How much do you want to play again with the {robot, screen}?	C [27]
How friendly is the {robot, screen}?	C [27]
Is the {robot, screen} exciting?	PE [28]
Is the {robot, screen} fun?	PE [28]
Does the {robot, screen} keep you happy during the game?	PE [28]
Is the {robot, screen} boring? (inverted)	PE [29]
Is playing with the {robot, screen} easy?	PEU [30]
Is communicating with the {robot, screen} easy?	PEU [30]
Is the {robot, screen} useful when playing the game?	PEU [30]
Is the {robot, screen} helpful when playing the game?	PEU [30]
Is playing with the {robot, screen} hard? (inverted)	PEU [30]

successfully completed the study and were provided with compensation for their time. This study was approved by the University of Southern California Institutional Review Board under protocol #UP-19-00185.

F. Measures

The participant preference of the embodiment (robot vs. screen) was measured using a three-factor five-point Likert scale, with questions from scales validated in previous works [27]–[30]. The three factors were: perceived enjoyment, companionship, and perceived ease of use.

The participants' engagement was quantified using identical criteria as in Clabaugh et al. [15], which also measures engagement in a game-based interaction. The participant was labelled as engaged if they responded to the robot's question, thought about the correct answer to the question, had a positive facial expression, and was looking at the robot as seen in the auditory and visual data captured by the camera. We represented the level of engagement as a binary variable (engaged/not engaged), as measured by a trained annotator. To ensure consistency, a secondary annotator independently annotated 10% of the videos selected at random. We measured inter-rater reliability with Cohen's Kappa, and achieved substantial agreement of k = .73, corresponding to an agreement on 86% of videos, similar to other works in engagement [15], [31].

IV. STUDY RESULTS

A. User Preference

Embodiment preference was determined by the difference in ratings between corresponding questions for the robot and screen conditions. The specific questions are shown in Table I. The combined responses for all factors are shown in Figure 3. We found high internal consistency for all factors: Perceived Enjoyment ($\alpha = .94$), Companionship ($\alpha = .91$), and Ease of Use ($\alpha = .89$). We evaluated significance with a Wilcoxon Signed-Rank Test and found a significant preference for the robot over the screen in factors measuring Companionship (Z = 9.0, p = .026) and Perceived Enjoyment (Z = 23.5, p = .018). We found no significant differences in Ease of Use (Z = 52.0, p = .399) and attribute this to the fact that both embodiments used the same vision system, which suffered from perceptual errors



Fig. 3: Participant responses to Likert-scale questions, grouped by measured construct.

(such as failing to detect the participant's off-camera thumb angle) about 20% of the time. We additionally note that many of the responses showed no preference for either the robot or screen due to the tendency of the participants to respond with similar values for all questions. The results therefore partially support H1, indicating that participants somewhat preferred to interact with the robot over the screen.

B. Choice Condition

In the final study block, participants could choose to play a game with the robot, play another game with the screen, or stop playing and end the session. One participant chose to play with the robot, five participants chose to play with the screen, and four participants chose to stop playing. While the sample size is too small to draw any conclusions, the confounding factors include the possibility that participants were too fatigued to continue playing or reluctant to require work of the experimenter to exchange the embodiments. Several participants expressed these sentiments while doffing the orthosis at the conclusion of the experiment.

C. Engagement

We found no significant differences in the participants' engagement between the two conditions, which does not support H2; our post-hoc analysis shows that there were individualized differences in how engagement changed in response to the robot's actions. We discuss those differences next in the context of personalizing the interaction.

V. MODELING ENGAGEMENT

We first explored whether there were differences across users in how engagement changed in response to the robot's actions. We modeled the evolution of engagement as a Markov chain, where engagement is a binary state variable that changes stochastically in discrete time-steps, after each of the robot's actions.

We define a transition matrix T that specifies how engagement $s \in S$ changes over time $T : S \times A \rightarrow \Pi(S)$. Since the change depends on the robot's action (*clarify, encourage* or *reward*), we parameterized the transition matrix by the robot's action $a \in A$.

A. Learning Personalized Models

To learn personalized engagement models, the first step is to represent how likely a participant is to become engaged or disengaged given a robot's action. We captured this with the transition matrix of the Markov chain. For each participant in the user study, we computed a transition matrix using



Fig. 4: Transition matrices of the two clusters found in the participant-based clustering method. Each matrix specifies the probability of becoming engaged (E) or disengaged (D) at the next time-step, given the current state.

maximum likelihood estimation from the sequence of the annotated engagement values.

We next explored whether participants cluster in terms of similar reactions to the robot's actions. Previous work [32] has shown that users can be grouped based on their preference on how to perform a collaborative task with a robot. We used a similar approach in the context of social interaction: we clustered participants from the study based how their engagement changed in response to the robot's actions.

We converted the transition matrices to vectors, then computed the distance between vectors using cosine similarity. We then performed hierarchical clustering [33], by iteratively merging the two most similar vectors into a cluster. The merged vector was formed by averaging the values of the two vectors. We selected the final number of clusters, so that each cluster contained at least two individuals. We transformed the vector of each cluster back to a transition matrix that specified how engagement changed for participants of that cluster.

We clustered participants at two different resolutions: 1) based on the user as a whole and 2) based on the users' response to each of the robot's three possible actions. The first clustering, based on each participant's holistic response to to robot actions, resulted in matching each participant to one cluster. We call this *participant-level clustering*. The second clustering, based on each participant's responses to each of the robot's action separately, required three different clustering iterations, one for each action, and resulted in having each participant matched to three clusters, one for each action. We call this *action-level clustering*.

B. Cluster Identification

Using the participant-level method, we found two main clusters, shown in Figure 4. In the first cluster, the *encourage* action had a greater likelihood of causing the participant's next state to be Engaged (E) than the *reward* action. In the second cluster we observed the opposite effect: the probability of changing from Disengaged (D) to Engaged (E) is lower for the *encourage* action than for the *reward* action. We observe that the *clarify* action has a small effect on changing a participant's engagement state in both clusters. Seven participants belonged to the first cluster, and three participants belonged to the second cluster. There were no clear factors that lead to the makeup of the participants in the clusters based on the background information collected in the study.

The action-level clustering method generated separate clusters for each action independently (Figure 5). Thus, a single participant is described as being a part of three actionlevel clusters. We observed three different types of matrices across the different actions:

- Type I indicates a high probability of the participant becoming Engaged, regardless of their previous state.
- Type II has an approximately equal probability of becoming Engaged or remaining Disengaged, if the participant was previously Disengaged.
- Type III features participants who are most likely to remain in the same state.

We found that the *clarify* action generated only Type II and III clusters, since most participants' engagement did not change based on that action, with three participants belonging to the Type II cluster and seven participants belonging to the Type III cluster when conditioning on the *clarify* action. The *reward* action generated Type I and Type II clusters, since most participants belonged to the Type I cluster and seven participants belonged to the Type I cluster and seven participants belonged to Type II cluster. This finding supports previous work [34] that showed positive reinforcement improving participants' engagement in computer-based animal guessing games.

Only the *encourage* action generated clusters of all three types. The *encourage* action had three participants that responded in alignment with the Type I cluster, five participants that aligned with the Type II cluster, and two participants that aligned with the Type III cluster.

We investigated whether the composition of the participants in each cluster was related to the demographic information we collected. Specifically, we analyzed cluster composition with a multinomial logistic regression of age, gender, and handedness onto cluster type and found no significant differences in composition between any clusters at either participant or action levels. Qualitatively, age was a minor component in the Encourage clusters; older ages appeared to be more associated with the Type II cluster (median age 16), whereas younger participants either fell into Type I (median age 14) or Type III (median age 10.5) clusters.



Fig. 5: Transition matrices of different clusters found in the action-based clustering method. Each matrix specifies the probability of becoming engaged (E) or disengaged (D) at the next time-step, given the current state.

C. Personalizing Robot Actions

If a robot knows the participant's cluster and adapts its actions to maximize engagement, to what extent does this improve the participant's engagement? Following prior work in simulating users based on personas [35], we show the benefit of using a personalized engagement model by modeling users based on the data from the user study. We focus on the action-level clustering method, since it generates different clusters for each robot action, resulting in a higher resolution model than the participant-level clusters using the same amount of data.

We modeled users based on each participant's transition matrices described in Section V-A. At each timestep, we have the user's current engagement state and the set of the ground-truth transition matrices for each action from the study. When the robot takes an action, the user's next engagement state is sampled from the probability distribution of the corresponding action and the current engagement state. This process is repeated for 100 timesteps, as determined by the number of turns in the in-person study. We additionally average our results over 100 runs for each participant to mitigate the random effects of the simulation and converge to the true mean of the engagement level over the course of the simulation.

Our strategy for selecting actions in the simulation is to maximize the likelihood of the user becoming engaged based on the estimated user clusters. For instance, if a user is said to be Type II for *clarify*, Type I for *encourage* and Type II for *reward*, and the participant is currently disengaged (D), then the robot would choose the encourage action, since the participant would become engaged (E) with probability 0.87 (compared to 0.48 for reward and 0.59 for clarify). These estimated clusters, however, are distinct from the true clusters used to simulate the users: for example, a user may truly be associated with the type II cluster for encourage actions, but we may erroneously select actions as if that user were associated with a type I cluster for the *encourage* action. We additionally imposed a 0.2 probability of the robot taking a clarify action no matter what state the participant was in to account for incorrect or illegible gestures; similar to the rate that was observed in the in-person study.



Fig. 6: Percentage of time that modeled users were engaged for different methods of robot action selection. Selecting actions based on the correct user clusters (a) keeps users more engaged, however selecting actions on incorrect user models (b) has an adverse effect. Considering the users as one group (c) performs similarly to the random baseline.

We computed two baselines: 1) the robot selects actions randomly, choosing *encourage* or *reward* action with equal probability; and 2) the robot selects actions to maximize engagement based on the transition matrix computed from the maximum likelihood estimate of all the participants. The second baseline, which we call the "impersonal strategy", is equivalent to having one cluster for every action; it does not account for individual differences.

Fig. 6 shows the average time the modeled users spent being Engaged in the activity for each condition. A twotailed paired t-test showed that modeled users spent significantly more time being Engaged when the robot selected actions given their cluster compared to taking random actions (t(11) = 9.370, p < .001). However, if the robot had a flawed model of the user, the modeled users were significantly less engaged over the course of the interaction compared to randomly selecting actions (t(131) = -4.692, p < .001). This result highlights the trade-offs that personalization may bring, especially in low-data scenarios.

The second baseline treated users as coming from one cluster. The personalized strategy significantly outperformed the impersonal strategy (t(11) = 4.984, p < .001). Furthermore, we cannot say that the impersonal strategy performed any differently than randomly selecting actions (t(11) = -.656, p = .525). This shows the importance of algorithmic design in interaction, and how incorrect assumptions can lead to data-driven models that are ineffective.

VI. DESIGN INSIGHTS

From the study conducted with a participants with CP, we developed the following design insights that may be useful when designing personalized algorithms for end-users.

The distribution of interaction preferences is often unbalanced. The clusters formed from this study revealed skews in the number of participants, most commonly with the majority cluster being composed of seven of the ten participants. This highlights the importance of understanding the lower-probability modes in which users interact with an implemented system through the use of qualitative tools such as user personas [35]. These are especially helpful when there are no apparent differences between the clusters.

Clustering user responses helps to reduce the design space. Our clusters reveal three main types of responses to each action. Users found each action as either highly engaging (Type I), sometimes engaging (Type II), or having no effect on engagement (Type III). Interestingly, we did not see actions that were highly disengaging or caused the participant to switch states with high probability. This indicates that those types of interaction are less common, and therefore less focus can be placed on considering how those cases would affect the interaction.

Certainty in user cluster or persona is critical in personalized algorithms. Our results show that misclassification of a user's cluster or persona leads to lower engagement than random selection. To build effective systems, it is critical to be certain that the inferred user's classification is correct. A system that personalizes to a given user should consider the risk of misclassification in selecting the level of certainty that is required to make decisions in an interactive scenario.

VII. LIMITATIONS AND CONCLUSION

Our results are limited by our small sample size, resulting from the practical challenges of recruiting participants with CP. Applying our clustering approach with more participants will likely result in more nuanced representations of user engagement levels. Our method also carries the inherent limitations of Markov-based models: it does not account for effects of the history of interaction, such as fatigue, or aspects of the interaction that are not modeled, such as participants getting distracted by other events in their environment.

Additionally, our user models are based on the assumption that users can be associated with a previously known classification; they do not account for new, previously unseen classifications. Implementation in a real-world setting would also require correct inference of the user's engagement level in real time, as well as accurate identification of the user's classification. In fact, our user models show that incorrect inference results in worse performance than random robot action selection.

This work introduces socially assistive robotics to the context of communicative gesture practice for users with cerebral palsy. Our user study shows that participants with CP preferred to interact with a socially assistive robot compared to a screen-based agent. While we did not observe significant differences in user engagement overall, our posthoc analysis showed that there are nuanced differences between modes of participants: some participants became more engaged after the robot gave encouraging feedback, while others responded better to rewarding feedback. We show that understanding how a user will react to different robot actions can be leveraged to design a more engaging experience.

REFERENCES

- M. Oskoui, F. Coutinho, J. Dykeman, N. Jette, and T. Pringsheim, "An update on the prevalence of cerebral palsy: a systematic review and meta-analysis," *Developmental Medicine & Child Neurology*, vol. 55, no. 6, pp. 509–519, 2013.
- [2] S. Winter, A. Autry, C. Boyle, and M. Yeargin-Allsopp, "Trends in the prevalence of cerebral palsy in a population-based study," *Pediatrics*, vol. 110, no. 6, pp. 1220–1225, 2002.
- [3] T. D. Sanger, "Toward a definition of childhood dystonia," *Current opinion in pediatrics*, vol. 16, no. 6, pp. 623–627, 2004.
- [4] M. J. C. Hidecker, N. Paneth, P. L. Rosenbaum, R. D. Kent, J. Lillie, J. B. Eulenberg, K. CHESTER, JR, B. Johnson, L. Michalsen, M. Evatt *et al.*, "Developing and validating the communication function classification system for individuals with cerebral palsy," *Developmental Medicine & Child Neurology*, vol. 53, no. 8, pp. 704–710, 2011.
- [5] J. Realmuto and T. Sanger, "A robotic forearm orthosis using soft fabric-based helical actuators," in 2019 2nd IEEE International Conference on Soft Robotics (RoboSoft). IEEE, 2019, pp. 591–596.
- [6] J. A. Buitrago, A. M. Bolaños, and E. Caicedo Bravo, "A motor learning therapeutic intervention for a child with cerebral palsy through a social assistive robot," *Disability and Rehabilitation: Assistive Technology*, vol. 15, no. 3, pp. 357–362, 2020.
- [7] D. Feil-Seifer and M. J. Mataric, "Defining socially assistive robotics," in 9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005. IEEE, 2005, pp. 465–468.
- [8] M. J. Matarić and B. Scassellati, "Socially assistive robotics," in Springer handbook of robotics. Springer, 2016, pp. 1973–1994.
- [9] E. Deng, B. Mutlu, M. J. Mataric *et al.*, "Embodiment in socially interactive robots," *Foundations and Trends*® *in Robotics*, vol. 7, no. 4, pp. 251–356, 2019.
- [10] C. Oertel, G. Castellano, M. Chetouani, J. Nasir, M. Obaid, C. Pelachaud, and C. Peters, "Engagement in humanagent interaction: An overview," *Frontiers in Robotics* and AI, vol. 7, p. 92, 2020. [Online]. Available: https://www.frontiersin.org/article/10.3389/frobt.2020.00092
- [11] O. Celiktutan, E. Sariyanidi, and H. Gunes, "Computational analysis of affect, personality, and engagement in human–robot interactions," in *Computer Vision for Assistive Healthcare*. Elsevier, 2018, pp. 283–318.
- [12] S. Jain, B. Thiagarajan, Z. Shi, C. Clabaugh, and M. J. Matarić, "Modeling engagement in long-term, in-home socially assistive robot interventions for children with autism spectrum disorders," *Science Robotics*, vol. 5, no. 39, 2020.
- [13] O. Rudovic, J. Lee, L. Mascarell-Maricic, B. W. Schuller, and R. W. Picard, "Measuring engagement in robot-assisted autism therapy: A cross-cultural study," *Frontiers in Robotics and AI*, vol. 4, p. 36, 2017.
- [14] O. Rudovic, H. W. Park, J. Busche, B. Schuller, C. Breazeal, and R. W. Picard, "Personalized estimation of engagement from videos using active learning with deep reinforcement learning," in 2019 IEEE/CVF

Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, 2019, pp. 217–226.

- [15] C. E. Clabaugh, K. Mahajan, S. Jain, R. Pakkar, D. Becerra, Z. Shi, E. Deng, R. Lee, G. Ragusa, and M. Mataric, "Long-term personalization of an in-home socially assistive robot for children with autism spectrum disorders," *Frontiers in Robotics and AI*, vol. 6, p. 110, 2019.
- [16] S. Rossi, F. Ferland, and A. Tapus, "User profiling and behavioral adaptation for hri: A survey," *Pattern Recognition Letters*, vol. 99, pp. 3–12, 2017.
- [17] A. Tapus, C. Ţăpuş, and M. J. Matarić, "User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy," *Intelligent Service Robotics*, vol. 1, no. 2, pp. 169–183, 2008.
- [18] D. Leyzberg, S. Spaulding, and B. Scassellati, "Personalizing robot tutors to individuals' learning differences," in 2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI). IEEE, 2014, pp. 423–430.
- [19] N. A. Malik, F. A. Hanapiah, R. A. A. Rahman, and H. Yussof, "Emergence of socially assistive robotics in rehabilitation for children with cerebral palsy: A review," *International Journal of Advanced Robotic Systems*, vol. 13, no. 3, p. 135, 2016.
- [20] A. Brisben, C. Safos, A. Lockerd, J. Vice, and C. Lathan, "The cosmobot system: Evaluating its usability in therapy sessions with children diagnosed with cerebral palsy," *Retrieved on*, vol. 3, no. 25, p. 13, 2005.
- [21] N. A. Malik, H. Yussof, F. A. Hanapiah, and S. J. Anne, "Human robot interaction (hri) between a humanoid robot and children with cerebral palsy: experimental framework and measure of engagement," in 2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES). IEEE, 2014, pp. 430–435.
- [22] "Qtrobot: Humanoid social robot for research and teaching," 2020. [Online]. Available: http://luxai.com/qtrobot-for-research/
- [23] E. Short, D. Short, Y. Fu, and M. J. Mataric, "Sprite: Stewart platform robot for interactive tabletop engagement. department of computer science, university of southern california," *Tech Report*, 2017.
- [24] C. Long and J. Kridner, "Meet beagle: Open source computing," 2019. [Online]. Available: https://beagleboard.org/
- [25] Stanford Artificial Intelligence Laboratory et al., "Robotic operating system." [Online]. Available: https://www.ros.org
- [26] M. V. Johnston and H. Hagberg, "Sex and the pathogenesis of cerebral palsy," *Developmental Medicine & Child Neurology*, vol. 49, no. 1, pp. 74–78, 2007.
- [27] K. M. Lee, N. Park, and H. Song, "Can a robot be perceived as a developing creature? effects of a robot's long-term cognitive developments on its social presence and people's social responses toward it," *Human communication research*, vol. 31, no. 4, pp. 538– 563, 2005.
- [28] J.-W. Moon and Y.-G. Kim, "Extending the tam for a world-wide-web context," *Information & management*, vol. 38, no. 4, pp. 217–230, 2001.
- [29] M. Heerink, B. Krose, V. Evers, and B. Wielinga, "Measuring acceptance of an assistive social robot: a suggested toolkit," in *RO-MAN* 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication, 2009, pp. 528–533.
- [30] V. Venkatesh, "Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model," *Information systems research*, vol. 11, no. 4, pp. 342–365, 2000.
- [31] O. Rudovic, J. Lee, M. Dai, B. Schuller, and R. W. Picard, "Personalized machine learning for robot perception of affect and engagement in autism therapy," *Science Robotics*, vol. 3, no. 19, 2018.
- [32] S. Nikolaidis, R. Ramakrishnan, K. Gu, and J. Shah, "Efficient model learning from joint-action demonstrations for human-robot collaborative tasks," in 2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI). IEEE, 2015, pp. 189–196.
- [33] D. Müllner, "Modern hierarchical, agglomerative clustering algorithms," arXiv preprint arXiv:1109.2378, 2011.
- [34] B. J. Fogg and C. Nass, "Silicon sycophants: the effects of computers that flatter," *International journal of human-computer studies*, vol. 46, no. 5, pp. 551–561, 1997.
- [35] A. Andriella, C. Torras, and G. Alenyà, "Learning robot policies using a high-level abstraction persona-behaviour simulator," in 2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 2019, pp. 1–8.